A contribution to the search for CP Violation in Λ0b four-body decays

Maxime Vernet

UFR Sciences et Technologies

Laboratoire de Physique Corpusculaire de Clermont-Ferrand

MASTER DEGREE IN SCIENCE OF MATTER
SECOND YEAR

SPECIALTY: **Particle Physics**

INTERNSHIP REPORT

A contribution to the search for CP Violation in Λ_b^0 four-body decays

Maxime VERNET

Supervisor: **Stéphane MONTEIL**

June 2015

Blaise Pascal University – IN2P3 – CNRS 4 avenue Blaise Pascal BP80026 63177 Aubière Cedex

Acknowledgment

I would first of all like to thank the Laboratoire de Physique Corupusculaire of Clermont-Ferrand and especially the LHCb team, for making me feel very welcome during these few months of internship. I would like to express my gratitude to my internship supervisor Stéphane Monteil, who offers fundamental knowledge and brings my prior wish to work in the field of particle physics to eagerness. Eventually, I would like to thank Jan Maratas, whom I worked with, particularly for his precious help in the tools for comprehension and for his systematical availability.

Contents

Introduction **1**

I Scientific Context

1 CP Violation in the Standard Model **2**
1.1 Overview . . . 2
1.2 CKM Matrix and CP Violation . . . 3
1.2.1 CKM Matrix . . . 3
1.2.2 CP Violation . . . 3

2 LHC and LHCb **5**
2.1 The Large Hadron Collider . . . 5
2.2 LHCb spectrometer . . . 6
2.2.1 VELO: Vertex Locator . . . 7
2.2.2 RICH: Ring Imaging Cherencov counter . . . 7
2.2.3 Tracking system . . . 7
2.2.4 Magnet . . . 7
2.2.5 Calorimetric system . . . 7
2.2.6 Muon System . . . 7
2.2.7 Particle IDentification system . . . 7

II Analysis

3 Objective **8**

4 Selection of charmless decays **9**
4.1 Principles of Boosted Decision Tree . . . 9
4.2 Discriminating variables . . . 9
4.3 Preselection . . . 10
4.4 Performance . . . 10
4.5 Agreement test between charmless distributions from Monte-Carlo and from real data . . . 12

5 Selection of charmed decays **14**
5.1 Identification of the decay mode . . . 14
5.2 Particle IDentification . . . 14
5.3 Invariant mass distributions . . . 15

6 Fit model **18**
6.1 Signal . . . 18
6.2 Partially reconstructed background . . . 19
6.3 Combinatorial backgound . . . 19
6.4 Fit function . . . 20

7 Fit Results **21**
7.1 Fit parameters . . . 21
7.2 Yields per channel . . . 21
7.3 Invariant mass distributions and fit results displays . . . 22

8 Estimation of B physics background yield present in the data **27**
8.1 Estimation in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ spectrum . . . 27
8.2 Estimation in the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ spectrum . . . 29

Conclusion **31**

III APPENDICES

A Boosted Decision Tree (BDT) **33**
A.1 Training . . . 33
A.2 Test . . . 33
A.3 Overtraining . . . 34

B Chebychev polynomials **35**

C B physics background estimation with other cut on PID variable of the proton. **36**

Bibliography **39**

Introduction

The Standard Model (SM) describes strong, weak and electromagnetic elementary interactions. It has been developed in the last 50 years in an interplay of experimental and theoretical achivements. Electroweak precision measurements allowed to predict in particular, the mass of the fundamental scalar of the model, the so-called Brout-Englert-Higgs bosonl[1]. The SM received thus an oustanding confirmation though precise measurements of the properties of the observed narrow bosonic resonance have still to be performed. There are however fundamental questions which are not answered in the SM. Their review is out of the scope of this report but one of them defines the context of my Master internship.
It seems that the SM fails at explaining the observed baryonic asymmetry in the universe. The search for new CP-violating phases, beyond the SM CKM one, is one of the main objectives of the LHCb experiment. If CP violation phenomena have been observed in kaon and B meson systems, CP violation in baryon decays has not been seen so far.
The 4-body charmless decays (involving $b \to u$ transitions) of b-baryons as Λ_b^0 and Ξ_b^0 could be the right place to measure, for the very first time, CP-asymmetries in baryon decays. The measurement of this CP-asymmetry requires however to master detection efficiencies and production asymmetries. It is necessary to use control channels to master empirically these experimental induced asymmetries mimicking CP-violation effects. The subject of my internship is in particular to study such control channels as $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ or $\Lambda_b^0 \to D^0 p \pi^-$.
The first part of this document gives a theoritical and experimental overview of the context in which CP violation measurements take place. The analysis presented in the second part consists of selecting baryon decays by using a multivariate analysis and building a fit model in order to well characterize control channels. Eventually, we will give an estimation of the B meson contribution in term of background in the sample.

Part I

Scientific Context

Chapter 1

CP Violation in the Standard Model

I am presenting in this section the necessary elements of the SM to introduce my master subject.

1.1 Overview

The Standard Model describes the known elementary particles (and their antiparticles) present in the Nature as well as the strong and electroweak interactions ([2], [3], [4]). The weak interaction requires to rank the quarks and leptons in doublets of weak isospin. The Table 1.1 orders them in three generations, doublets of $SU(2)_L$.

Fermions	Generations		
	1^{st}	2^{nd}	3^{rd}
Quarks	up u	charm c	top t
	down d	strange s	beauty b
Leptons	Electron e	Muon μ	Tau τ
	ν_e	ν_μ	ν_τ

Table 1.1: Fundamental fermions

The Electroweak SM generates the weak and electromagnetic interactions from local invariance under $SU(2)_L \otimes U(1)_Y$ transformations. The electroweak bosons are the massive intermediate bosons Z^0 and $W^{\pm}$. The photon emerges "naturally" as the mediator of the electromagnetism. Eventually, the strong interaction is described from the $SU(3)_C$ symmetry. C stands for the charge of the interaction (colour) and gluons are the mediators of this interaction.

The quarks (up to the b quark) hadronize before their decay. The simplest strongly binded particles (hadrons) are:

- the mesons: combinaisons of one quark and one antiquark as $|\bar{q}q'\rangle$
- the baryons: combinaisons of three quarks as $|qq'q''\rangle$

In this work, we will study the weak Λ_b^0 decays in 4-body. The Λ_b^0 is described in the quark model as a state $|bud\rangle$

1.2 CKM Matrix and CP Violation

1.2.1 CKM Matrix

The Cabibbo-Kobayashi-Maskawa matrix describes the quark mixing observed in the charged currents of the weak interaction. It emerges in the SM after the spontaneous symmetry breaking when diagonalizing the quark mass matrices. It reads as :

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

This 3×3 matrix is complex and unitary. This unitarity brings 9 constraints. Moreover, each quark field comes with an arbitrary phase which can be redefined up to a single phase (global) bringing 5 more constraints. Hence, the CKM matrix can be described from 4 independant parameters, free parameters of the SM. A possible parameterization of the CKM matrix considers three Euler angles (θ_{12}, θ_{23},θ_{13}), and one phase δ_{13} which is the unique source of CP violation in the Standard Model:

$$V_{CKM} = R_{23}(\theta_{23}, 0) \otimes R_{13}(\theta_{13}, \delta_{13}) \otimes R_{12}(\theta_{12}, 0).$$

In this way we have:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}.$$

with $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$.
The experimental hierachy observed in the CKM elements magnitudes provides a guideline for another relevant parameterization, introduced by L. Wolfenstein [5]. In this parameterization, the four parameters are λ, A, ρ and η and the probability to observe a diagonal flavour mixing between two quarks are related to the power of the parameter λ, e.g. $V_{ub} \propto \lambda^3$ for the $b \to u$ transition.

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

We will in particular discuss in this report the decay $\Lambda_b^0 \to \Lambda_c^+ \pi^-$. This process is governed by a $b \to c$ transition hence involving the magnitude of the matrix element $|V_{cb}|$.

1.2.2 CP Violation

The symmetry C (for Charge conjugation) changes particles in antiparticles (and vice-versa). The P symmetry (Parity) is the space reflexion as $x \to -x$. The P violation was suggested for the very first time with the $(\theta - \tau)$ puzzle and confirmed in 1956 by the Wu's experiment [6]. Actually we know that only left-handed[1] particles and right-handed[1] antiparticles are selected in

[1] Left(Right)-handed refers here to chirality. When fermion has no mass, it is equivalent to helicity. Left(Right)-handed means that the impulsion $\vec{p}$ is (anti-)aligned with the spin $\vec{s}$

process including charged current weak interaction. In other words, it means that C-symmetry and P-symmetry are (maximally) violated by the weak interaction.
What about the CP symmetry ?
The CP violation was first observed in the mixing of the neutral kaons K^0 and $\bar{K^0}$. In Quantum Mechanics we can express weak eigenstates as a linear combinaison of the mass eigenstates. Then we can apply CP symmetry as:

$$\begin{cases} CP\left|K_S^0\right\rangle = \frac{1}{\sqrt{2}}\left(CP\left|K^0\right\rangle + CP\left|\bar{K^0}\right\rangle\right) = +\left|K_S^0\right\rangle \ , \\ \\ CP\left|K_L^0\right\rangle = \frac{1}{\sqrt{2}}\left(CP\left|K^0\right\rangle - CP\left|\bar{K^0}\right\rangle\right) = -\left|K_L^0\right\rangle \ . \end{cases}$$

It means that if CP were a conserved quantity, we should have:

$$\Rightarrow \left|K_S^0\right\rangle \rightarrow \pi^+\pi^- \ ; \ \eta_{CP} = +1 \ ,$$

$$\Rightarrow \left|K_L^0\right\rangle \rightarrow \pi^+\pi^-\pi^0 \ ; \ \eta_{CP} = -1 \ .$$

in order to conserve CP-eigenvalue.
J.H. Christenson and J. W. Cronin [7] observe in 1964 that the K_L^0 decays into two pions:

$$|\eta_{+-}| = \frac{A(K_L^0 \rightarrow \pi\pi)}{A(K_S^0 \rightarrow \pi\pi)} = (2.232 \pm 0.011) \times 10^{-3} [8] \ .$$

which means that CP symmetry is slightly violated.

Beyond this first observation, CP violation was measured in the decays of kaons (2001), the interference between mixing and decays of B^0 mesons (2001), decays of B mesons (2004) and decays of B_s mesons (2013). This short list suggests that this phenomenon is difficult to study in reason of tiny asymmetries. The abundant production at LHC of b-baryons makes the LHCb experiment the good laboratory to find for the first time the CP violation in baryon decays.
Still, all CP measurement from mesons are very consistenly described within the SM. The baryons are a new territory to explore and try to find deviations to the SM prediction.

Chapter 2

LHC and LHCb

2.1 The Large Hadron Collider

The LHC is the largest proton-proton collider ever built. With a circonference of 27 km and an energy in the center of mass around 13 TeV (for 2015), the LHC is the machine nowadays to test and actually invalidate the Standard Model. The Standard Model received however an oustanding confirmation, with the discovery in 2012 of a narrow bosonic state, very much similar in its properties to the fundamental BEH scalar of the Standard Model[9].
There are four main collision points at LHC where are installed four detectors: ALICE, ATLAS, CMS and LHCb.

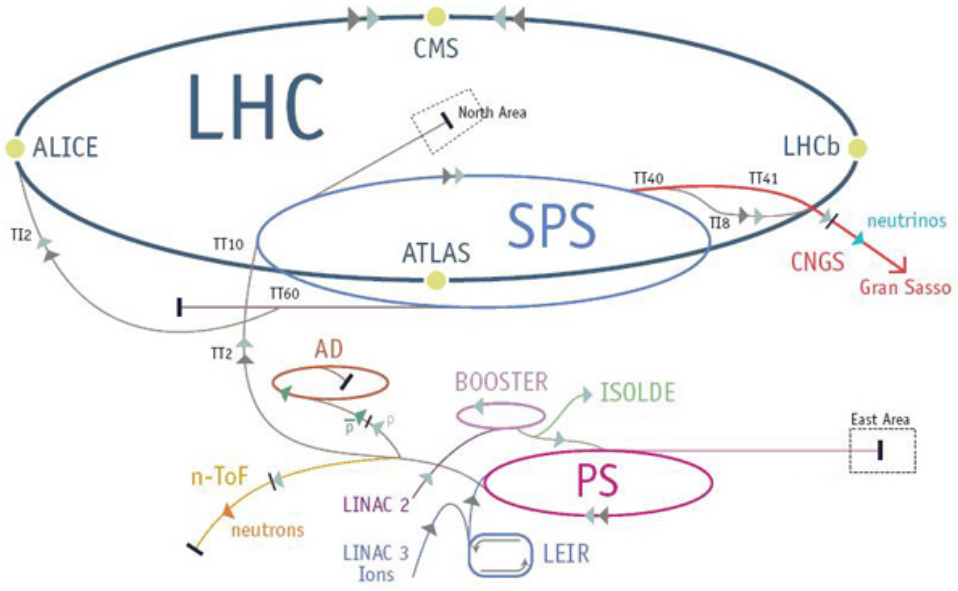


Figure 2.1: Representation of the LHC acceleration chain

My internship is taking place in the LHCb collaboration, operating the LHCb spectrometer at LHC. This experiment is dedicated to the study of rare decays of heavy flavours (b,c,τ) and the search for new CP-violating phases.

2.2 LHCb spectrometer

LHCb is one of the main experiments at the LHC. In this collider, there is a large $b\bar{b}$ pair production ($\sigma_{b\bar{b}} = 500\ \mu$b, which means typically 10^{12} per year) and all b-hadrons are produced: for our purpose, the hadronization fraction ($b \to \Lambda_b^0$) is 0.30 ± 0.03 [8].
$b\bar{b}$ pairs are predominantly produced in the same backward or forward cone: the LHCb geometry is hence defined as a single-arm spectrometer. The opposite phase space region is not instrumented as a trade-off between geometry, cost and luminosity arguments. The spectrometer overview is shown on the figure 2.2.
The detector is actually made of different sub-detectors wich manage to get as much information as possible during collisions. We will in the following briefly describe these sub-detectors as they are reported on the figure 2.2, form the left to the right.

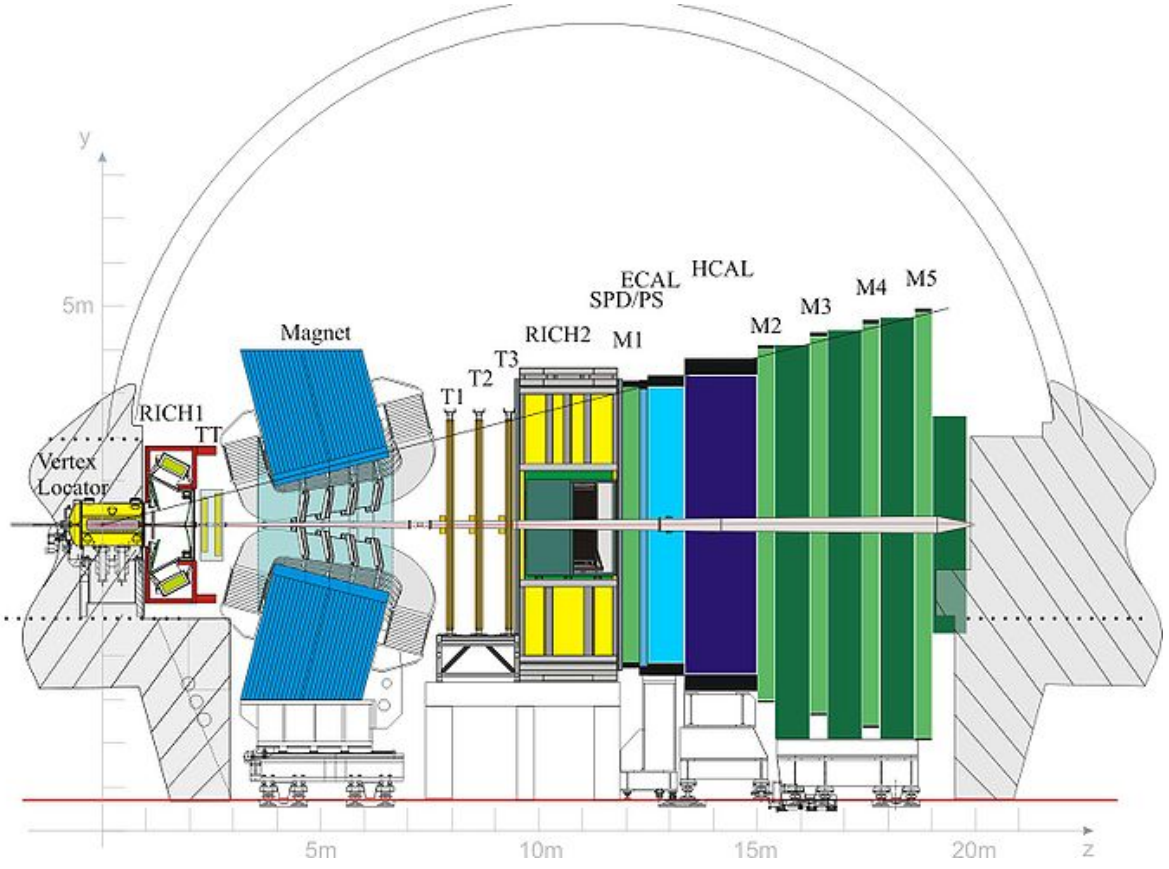


Figure 2.2: Transverse view of the LHCb detector

2.2.1 VELO: Vertex Locator

The Vertex Locator (VELO) is the detector closest to the interaction point. With a precise measurement of the r and ϕ coordinates[1] of charge particle track hits, this detector is capable to reconstruct secondary vertices generated by the decays of b-, c-hadrons or τ leptons. It also determines the primary vertex of the collision, where the paticles of interest are produced.

2.2.2 RICH: Ring Imaging Cherencov counter

Two Cherenkov detectors are used to perform an efficient particle identification. Indeed, the tracking system from LHCb experiment is not sufficient to get $\mathrm{d}E/\mathrm{d}x$ informations and then identify the particles. By using Cherenkov effect and by measuring the Cherenkov angle, it is possible to get a precise information of the particle's mass provided the momentum measured by the tracking system. RICH1 covers the low momentum range (≤ 60 GeV), and RICH2 covers the high one (up to 100 GeV).

2.2.3 Tracking system

There are four stations of tracker set upstream (TT) and downstream (T_1, T_2, T_3) the magnet. The reconstruction of the tracks is made by matching the TT hit with the (T_1, T_2, T_3) hits with an helix going through the not-instrumented magnet.
The TT is a silicon-based tracker. The stations T_1, T_2, T_3 have two different technologies in the same apparatus: silicon for the inner part to cope with the multiplicity and gaz straws for the outer part.

2.2.4 Magnet

The dipole magnet used in the LHCb experiment is a warm magnet which delivers an integrated magnetic field of 4 T.m. Its polarity can be reversed. This instrument plays thus a central role in the CP asymmetries measurement to control the charge asymmetries.

2.2.5 Calorimetric system

The calorimetric system is made of an hadronic calorimeter and an electromagnetic one. Through a destructive way, calorimeters give a measurement of the energy of the particles. The LPC LHCb team has the responsability of the Preshower of the electromagnetic calorimeter.

2.2.6 Muon System

As its name says, the muon system is dedicated to detect the presence of muons during an event. As for the calorimeter, it plays a leading role in the selection of the events (hardware trigger).

2.2.7 Particle IDentification system

The particle identification plays a central role in the understanding of the hadronic decay of b-hadrons event. With the two Cherenkov detectors, the momentum spectrum is mostly covered. During the selection of the decays of interest for the study of this report, we will see that PID variables are mandatory to reduce the $\pi \leftrightarrow K$ misidentification.

[1] Actually the z-coordinate is given by the modules which consitute the VELO, by considering the z-axis along the beam.

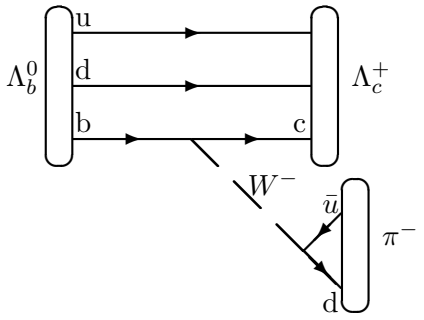
Part II

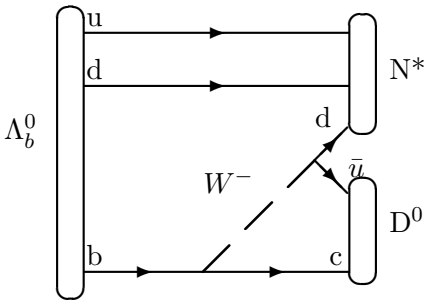
Analysis

Chapter 3

Objective

This work is developed in the context of searching for CP violation phenomena in baryons decays, by selecting four body charmless b-baryon decays $X_b \to phh'h''$, where h, h', h'' can be a pion or a kaon. Of particular interest is the decay $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$. It proceeds by two distinct topologies (tree and penguin) of same magnitude. Large interferences are then expected involving the weak phase. This interference pattern is further enhanced by the presence of rich structures at low invariant mass at threshold (e.g. $\Lambda^* \to pK$, $N^* \to p\pi$ ...). The very same final state can be reached through the decay $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^+\pi^-)\pi^-$ and can serve as a control channel, both for the control of CP asymmetries and mass fit model building. One of the aims of this analysis (described in section 5) is hence to fit invariant mass distributions of the Λ_b^0 ($m_{\Lambda_b^0} = 5619.4 \pm 0.7$ MeV)[8] decaying into two different charmed processes: via Λ_c^+ ($m_{\Lambda_c^+} = 2286.46 \pm 0.14$ MeV)[8] or via D^0 ($m_{D^0} = 1864.86 \pm 0.13$ MeV)[8], determined in advance. The decays of interest are represented on the diagrams below:





More generally, we are considering all the control channels given by different decays of Λ_c^+ and D^0:

(1) $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^-\pi^+)\pi^-$
(2) $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$
(3) $\Lambda_b^0 \to \Lambda_c^+(\to pK^-K^+)\pi^-$
(4) $\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$

The next chapter is dedicated to the selection of the events (both charmless and control channels). The following chapters will address the study of the control channels, in order to describe, in a second part, the b-hadron invariant mass fit procedure. The results of the fit procedure are given in the section 7.

Chapter 4

Selection of charmless decays

The purpose of this chapter is to describe the selection algorithm elaborated with the newly available data from a recent processing of the Run I data. The aim of this selection is to reject combinatorial background and preserve the signal decays. In the following, principles of performing Boosted Decision Tree are briefly describe[1]. Then, variables of discrimination and results are presented. Most of the tools were prepared in the framework of Jan Maratas' PhD. My contribution was to produce the multivariate variable for the newly available data as well as checking a posteriori the agreement between Data and Monte-Carlo for the individual discriminative variables.

4.1 Principles of Boosted Decision Tree

A Boosted Decision Tree (BDT) is a multivariate discriminant which takes into account the non linear correlations between input variables. Generally, signal events to train the BDT are fully simulated Monte-Carlo events through the entire detector simulation chain and are reconstructed with the same algorithms as data. Background events are taken directly from the data.
The signal events in our case are $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ charmless decays (MC). The background we are fighting against is combinatorics (association of unrelated tracks) taken from a mass region inconsistent with the signal.

4.2 Discriminating variables

By considering "X" as a candidate for Λ_b^0 and "h_i" the reconstructed tracks in an event, the variables used in order to separate Signal and Background are described in the following. All χ^2 estimators are defined with respect to the number of degrees of freedom.

1. $P_T(X_b^0)$: transverse momentum of the X_b^0 particle. The signal candidates have a harder momentum spectrum.

2. $\eta(X_b^0)$: pseudo-rapidity of the X_b^0 particle;

3. $\chi^2_{IP}(X_b^0)$: measure of the impact parameter of the X_b^0 particle. The impact parameter is the distance of closest approach of the particle trajectory to a reference point, here the primary vertex, location of the production of the b-baryon. It is expected that signal candidates Λ_b^0 have a small impact parameter.

[1] More details are given in Appendix A

4. $\chi^2_{FD}(X_b^0)$: measure of the flight distance of the X_b^0 particle. Λ_b^0 particle is decaying via weak interaction and will fly before decaying (typical flight distance is 1 cm in LHCb).

5. $\theta_{DIRA}(X_b^0)$: angle formed by the vertices (primary vertex-secondary vertex vector) and the momentum of the candidate.

6. $\chi^2_{EV}(X_b^0)$: represents the quality (expressed in χ^2/ndof) of the vertexing of the four candidate tracks. It is expected to be small for signal events.

7. $\Delta\chi^2_{EV}(X_b^0)$: measures the clearliness of the candidate vertex. Any additional track to the four candidate tracks vertex must degrade the vertexing quality for signal events.

8. $P_T^{asym}(X_b^0)$. Transverse momentum of all tracks contained in a cone formed around the X_b^0 momentum. This quantity measures the isolation of the decay of interest.

9. $\Sigma_i\chi^2_{IP,h_i}(h_i)$: sum over estimators of the impact parameters of daughters tracks h_i with respect to the primary vertex. The sum should be large if the tracks are coming from the X_b^0 decay.

4.3 Preselection

In order to prepare the training sample, we first apply a preselection. This selection consists of rejecting events which could not be of interest because of mandatory steps in the detection procedure which are not completed (e.g. trigger), or because of characteristics which directly eliminate the event as event of interest (e.g. one of the track coming from X_b^0 is a muon).

4.4 Performance

Multivariate analysis performances consist of having a good separation between signal and background events without any overtraining. The training sample, in the case of this analysis, is composed by Monte-Carlo simulation for the signal and by right-handed side-band data events for the background. Results given throughout the following figures are for the data of 2012.

Few comments are in order to describe these results :

- As we can see on the first figure 4.1, few variables are more discriminant than others. However, even if a certain variable is not very useful in term of seperation between background and signal, the global separation will be more efficient if we consider it, because of its correlation with other variables.

- The figure 4.2 testifies that the BDT is not overtrained. The training and the test sample responses are coherent, and an educated cut on the BDT variable (which carries the selection done by the BDT) can be set (around 0.3) to separate optimally the background from the signal.

- The global performance of the MVA selection is given on the figure 4.3. Excellent rejection of combinatorics is reached.

This selection is far more complex than what is reported here. However, it was an occasion for me to deal for the first time with a multivariate tool.

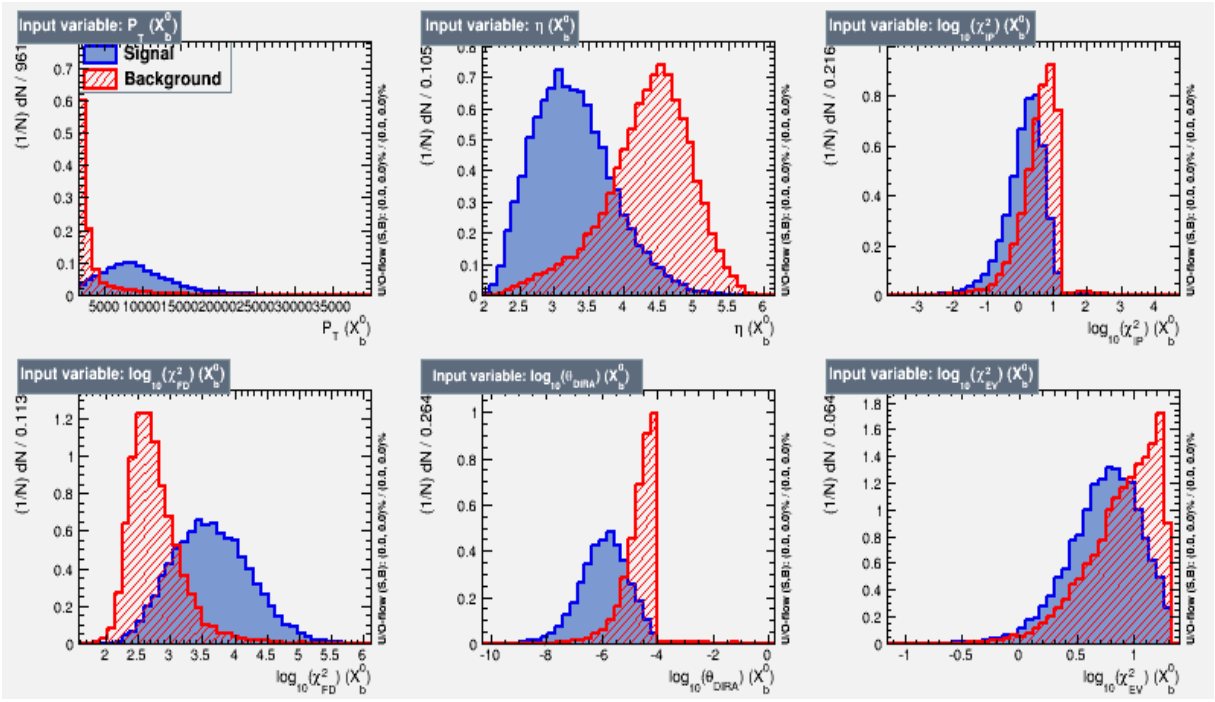


Figure 4.1: A selection of discriminating variables distributions (2012). Signal is in blue, background in red.

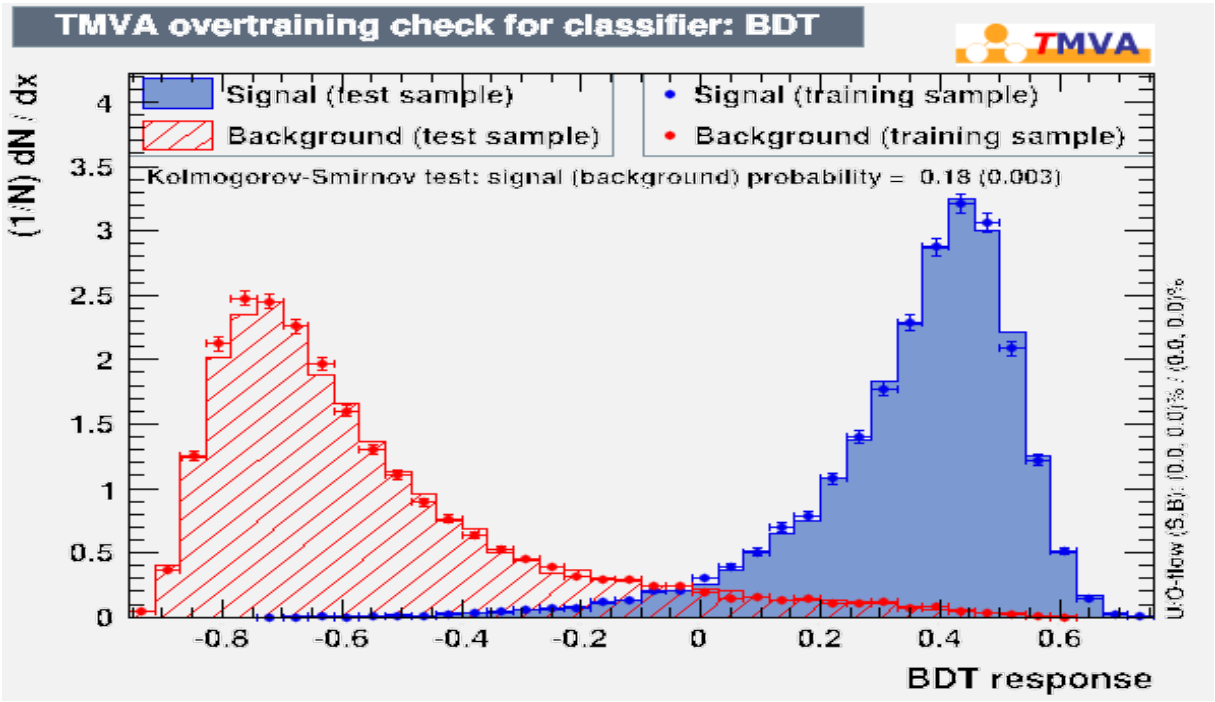


Figure 4.2: Overtrain plot for BDT classifier. (2012)

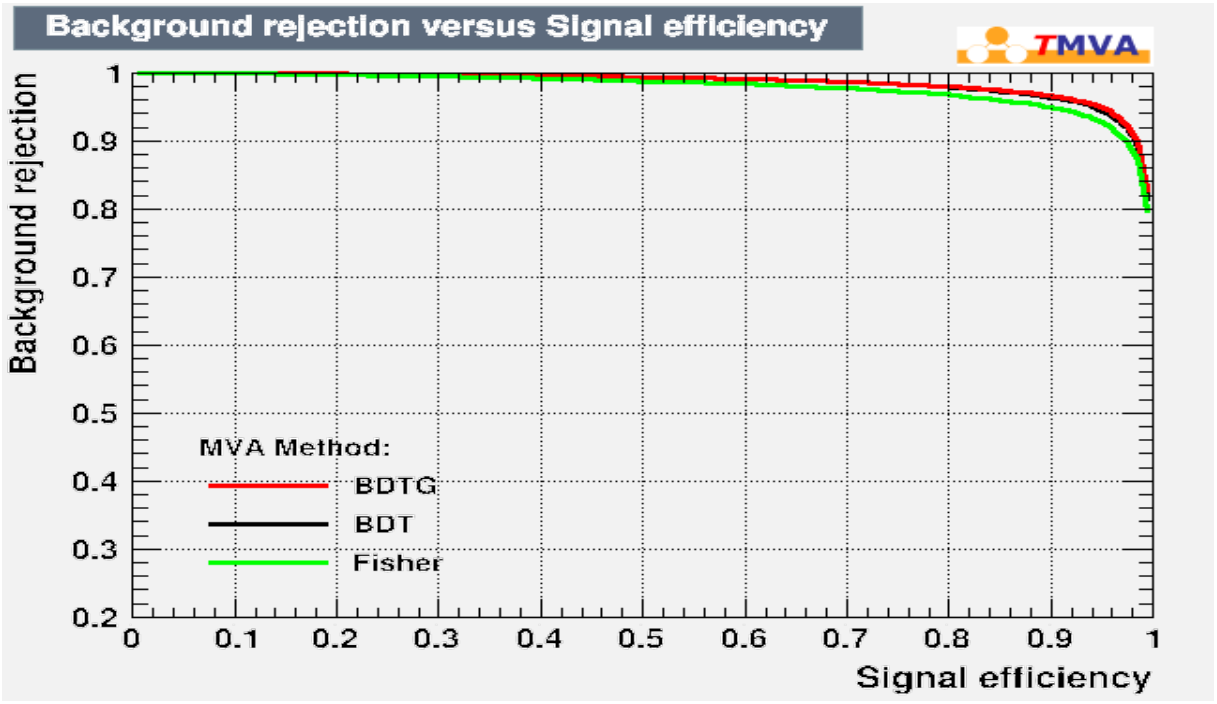


Figure 4.3: Receiver Operating Characteristic curves. Illustration with 2012 data. Three different MVA discriminations are compared. The BDT shows the best overall performance.

4.5 Agreement test between charmless distributions from Monte-Carlo and from real data

It is interesting to check the agreement between data and MC events for signal candidates. This can be seen as a measure of the optimality of the tool. This is done for events where:

$$m_{X_b^0} \in [m_{\Lambda_b^0} - 2\sigma; m_{\Lambda_b^0} + 2\sigma] \text{ with } \sigma \simeq 15 \text{ MeV}.$$

Most of the variables show a satisfactory agreement between data and Monte-Carlo. Some known discrepancies show up though.It has been however checked that the discrimination power is better with those variables in.

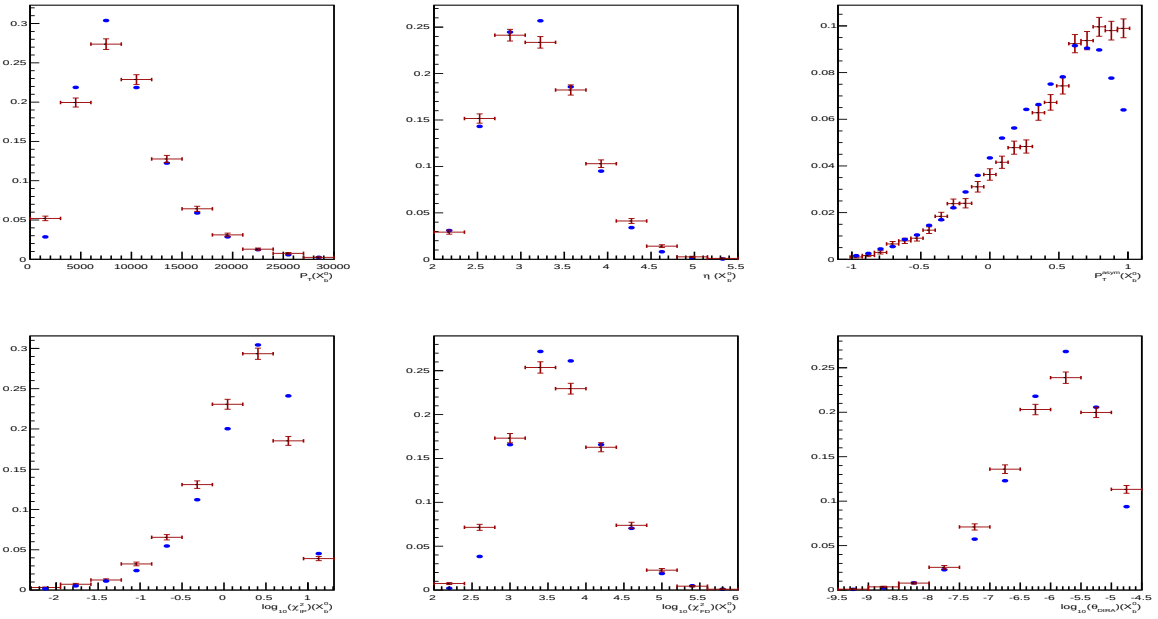


Figure 4.4: First part of the distributions of the discriminant variables. MC data in red, real data in blue.

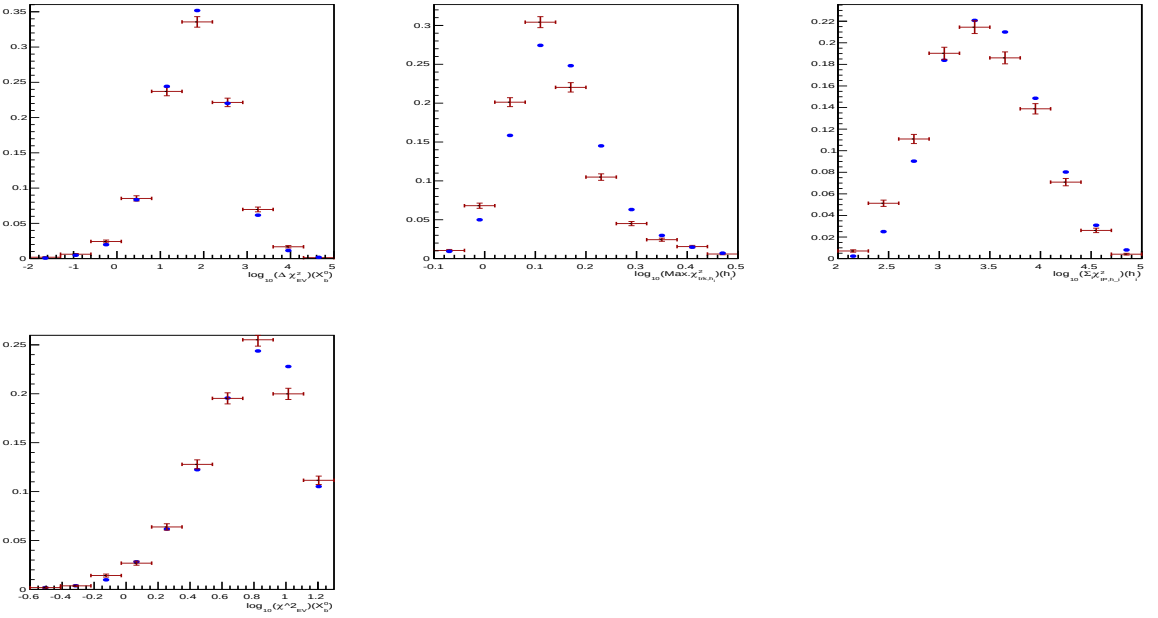


Figure 4.5: Second Part of the distributions of the discriminant variables. MC data in red, real data in blue.

Chapter 5

Selection of charmed decays

The data recorded by the LHCb experiment during LHC Run I have been analyzed prior to the beginning of my interniship. Let me summarize here the specific selection of the b-baryons we are interested in. Four-body charmless decays are selected on the basis of their vertexing properties only. Some of the final state particles are expected to have a soft momentum. In order to preserve their yields, no kinematical cut is applied.
In the case of this work, we are interested by selecting and characterizing the charmed decays. We want indeed to reconstruct the two decays $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ and $\Lambda_b^0 \to D^0 p \pi^-$ where $D^0 \to K^- \pi^+$ and $\Lambda_c^+ \to pK^-\pi^+$, pK^-K^+, $p\pi^+\pi^-$.
For the selection, only the events where $\mathrm{m}_{\Lambda_b^0} \in [5200,6200]$ MeV are considered, and a cut on the BDT variable is applied, as explained in the previous chapter.

5.1 Identification of the decay mode

In the events, four tracks are reconstructed ($h1^+h2^-h3^+h4^-$). The proton is identified as h1 by an a priori cut on the particle identification variable (PID). For each event, we reconstruct the invariant mass of three or two tracks[1] and if this invariant mass is contained in a mass window around the Λ_c^+ mass or the D^0 mass, the event is identified as a charmed Λ_b^0 decay.
For example, for the charmed decay $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ we proceed with the selection :

$$\Big(|m_{\Lambda_c^+} - M_{h1h2h3}| < 75 \text{ MeV}\Big) \text{ OR } \Big(|m_{\Lambda_c^+} - M_{h1h3h4}| < 75 \text{ MeV}\Big)$$

where M_{hihjhk} are the invariant mass reconstructed from the set of tracks $h_i h_j h_k$.

5.2 Particle IDentification

In order to strongly reduce the spectrum contamination linked to the misidentification between a kaon and a pion, we also impose a selection on the PID: for each track, we require a condition on the PID variables[2], which are related to the probability to consider the particle we are looking for: for example if we want the track h4 to be a kaon and not a pion, we should require:

$$(\text{probNN}(h2 = \pi) - 1))^2 + \text{probNN}(h2 = K)^2 > 0.55$$

[1] $\Lambda_c^+ \to$ 3 tracks and $D^0 \to$ 2 tracks
[2] denoted ProbNN

The left-handed quantity has to be lower than 0.55 in the reversed case. This condition has been optimized before my internship.
In this way, each event is selected in one decay mode only. Let's see an example of the selections to sum up this procedure: for the decay (4) $\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$ we build the data set with the following conditions:

$$\text{OR}\left\{\begin{array}{l} \text{AND}\left\{\begin{array}{l} |1864.86 - M_{h3h2}| < 75 \text{ MeV} \\ M_{h3h2} - 1864.86 < 30 \text{ MeV} \\ (\text{probNN}(h2 = \pi) - 1))^2 + \text{probNN}(h2 = K)^2 > 0.55 \\ (\text{probNN}(h3 = \pi) - 1))^2 + \text{probNN}(h3 = K)^2 < 0.55 \\ (\text{probNN}(h4 = \pi) - 1))^2 + \text{probNN}(h4 = K)^2 < 0.55 \end{array}\right\} \text{PID selection} \\ \\ \text{AND}\left\{\begin{array}{l} |1864.86 - M_{h3h4}| < 75 \text{ MeV} \\ M_{h3h4} - 1864.86 < 30 \text{ MeV} \\ (\text{probNN}(h2 = \pi) - 1))^2 + \text{probNN}(h2 = K)^2 < 0.55 \\ (\text{probNN}(h3 = \pi) - 1))^2 + \text{probNN}(h3 = K)^2 < 0.55 \\ (\text{probNN}(h4 = \pi) - 1))^2 + \text{probNN}(h4 = K)^2 > 0.55 \end{array}\right\} \text{PID selection} \end{array}\right.$$

One can notice here that an additional condition is set in this example. This asymmetric cut on the invariant mass is another way to limit cross-feeds between all spectra of interest.

5.3 Invariant mass distributions

The result of this selection procedure can be seen in figures 5.1 to 5.4 where invariant mass distributions for the Stripping 21 (2012) data sample are displayed for:

$\to$ $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^-\pi^+)\pi^-$

$\to$ $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$

$\to$ $\Lambda_b^0 \to \Lambda_c^+(\to pK^-K^+)\pi^-$

$\to$ $\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$, respectively

Several comments are in order to describe these distributions:

- the Λ_b^0 signal is unambiguously dominating the spectra for most of them;
- the combinatorial background, populating the right-handed side-band of the invariant mass distribution is modest, as a result of the Λ_c^+ and D^0 mass cuts but also because of the BDT cuts;
- the left-handed side-band is dominated by partially reconstructed Λ_b^0 decays, where a particle (a pion) has not been reconstructed (e.g. $\Lambda_b^0 \to pK^-\pi^+\pi^0\pi^-$).

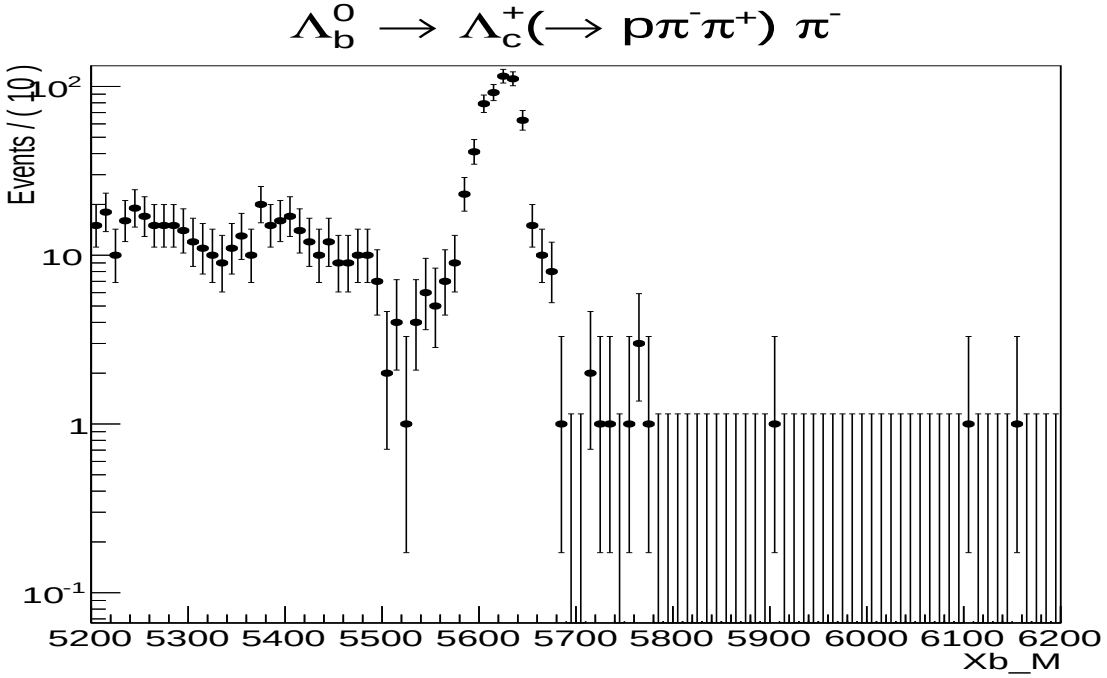


Figure 5.1: Invariant mass distribution for the decays $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^-\pi^+)\pi^-$, all selections applied, in logarithmic scale (data of 2012).

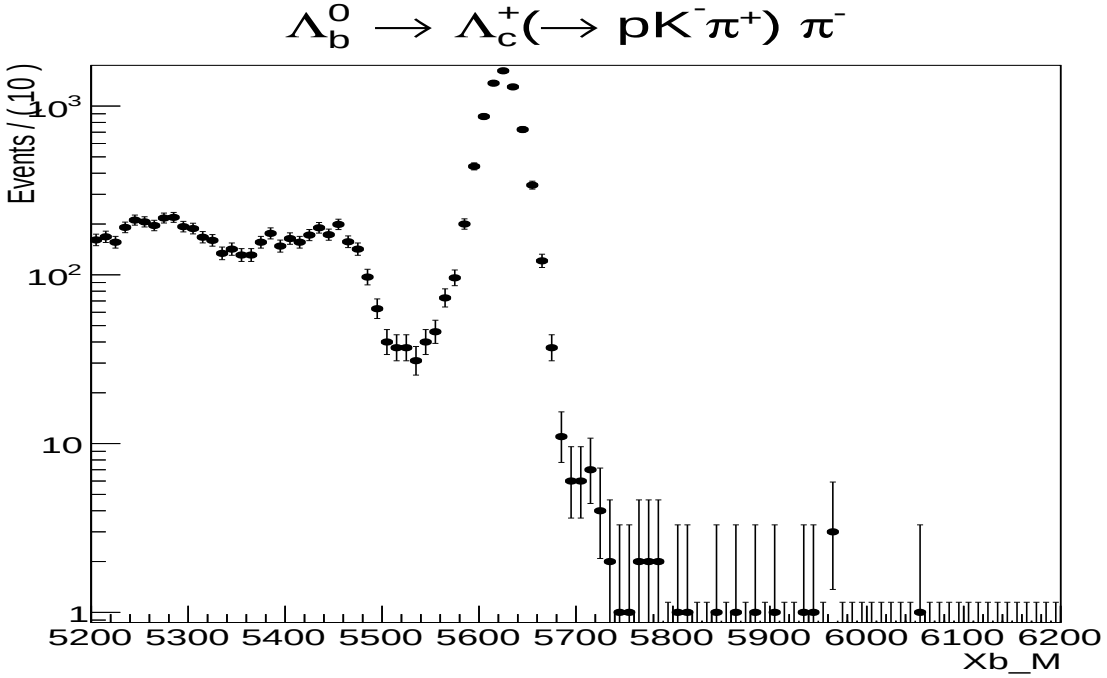


Figure 5.2: Invariant mass distribution for the decays $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$, all selections applied, in logarithmic scale (data of 2012).

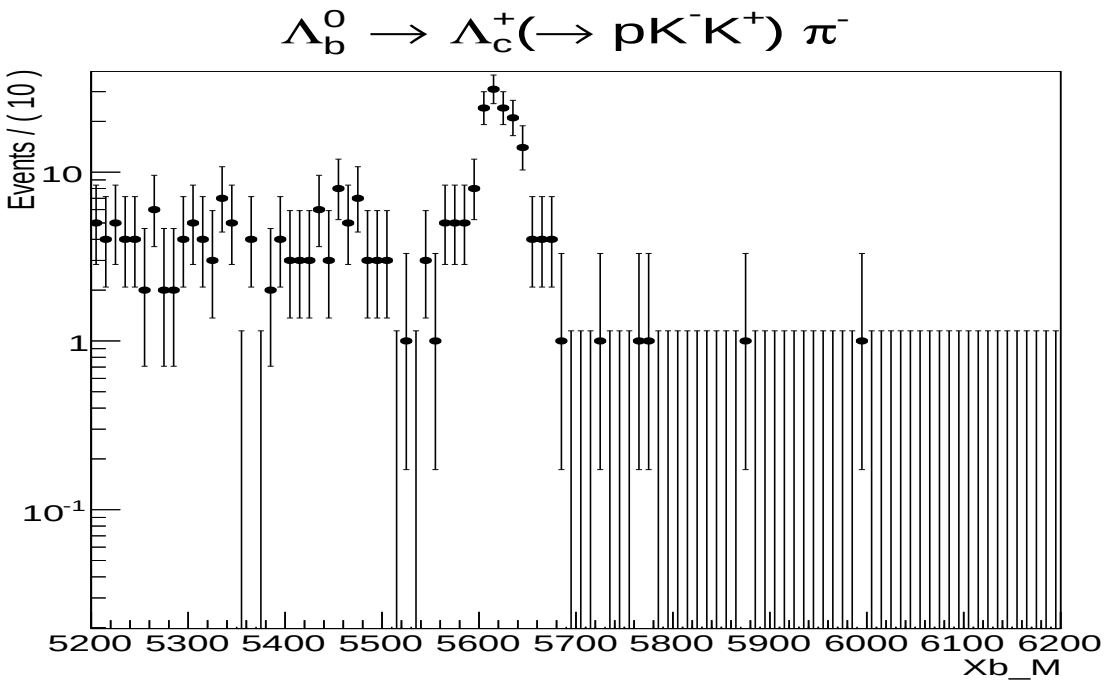


Figure 5.3: Invariant mass distribution for the decays $\Lambda_b^0 \to \Lambda_c^+(\to pK^-K^+)\pi^-$, all selections applied, in logarithmic scale (data of 2012).

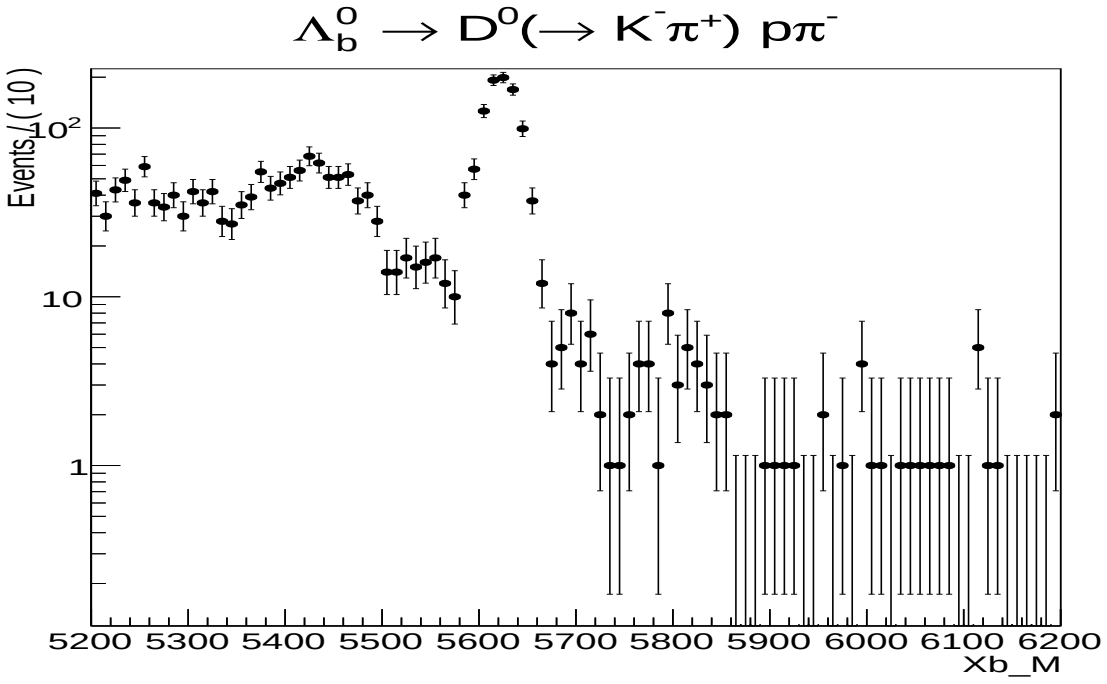


Figure 5.4: Invariant mass distribution for the decays $\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$, all selections applied, in logarithmic scale (data of 2012).

Chapter 6

Fit model

Following the initial study of the invariant mass distributions presented in the previous chapter, we build a probability density function which takes into account of signal, partially reconstructed background and combinatorial background. Well characterized spectra presented in section 5.3, which are $b \to c$ transitions, are indeed needed to determine later on a fit model for charmless ones, or educates some unknown shapes, such as partially reconstructed decays.

6.1 Signal

The invariant mass distribution of signal events is modelled with a double CrystalBall with asymmetric tails. The left-handed tail accounts for the radiative tail coming from Bremsstralhung effect. The right-handed one describes the imperfections of the tracking system. The core of the distribution is a Gaussian function. Its equation is given by [10]:

$$f_{\mathrm{CB}}(x) = \left.\frac{\left(\frac{n}{|\alpha|}\right)^n . e^{-\frac{\alpha^2}{2}}}{\left(\frac{n}{|\alpha|} - |\alpha| - x\right)^n}\right|_{x<-|\alpha|} \quad ; \quad \left. -\frac{1}{2}\left(\frac{x-\mu_{\Lambda_b^0}}{\sigma_{\Lambda_b^0}}\right)^2\right|_{x>-|\alpha} \tag{6.1}$$

$$\text{with} \begin{cases} \mu_{\Lambda_b^0}\text{: mean value} \\ \sigma_{\Lambda_b^0}\text{: width (core resolution)} \\ \alpha,\, n\text{: parameters of the tails of the function for } x < \mu_{\Lambda_b^0} \end{cases}$$

Full Monte-Carlo simulations of the signal decays are used to determine the parameters describing the tails as follow:

$$\frac{f_1}{f_0} = 0.406 \pm 0.10 \;\; ; \;\; \frac{\alpha_1}{\alpha_0} = -1.39 \pm 0.26 \;\; ; \;\; \frac{n_1}{n_0} = 1.72 \pm 0.17 \tag{6.2}$$

where f_i is the amount of signal related to the i^{th} CrystalBall function. Relations given in equation 6.2 are the ones used to determine parameters values which are used to perform the fit to the data spectrum.

6.2 Partially reconstructed background

The partially reconstructed background contains Λ_b^0 decays in 5-body with a track which is not reconstructed, and then, mimick the final state of interest, but not at the correct invariant mass. An Argus function convoluted with a Gaussian is used to model the partially reconstructed background. The probability density function of an Argus is given by [10]:

$$f_{\text{Argus}}(x) = x\left(1-\left(\frac{x}{m_0}\right)^2\right)^p \exp\left(c\left(1-\left(\frac{x}{m_0}\right)^2\right)\right) \quad \text{with} \quad \begin{cases} m_0\text{: threshold} \\ p\text{: power} \\ c\text{: slope} \end{cases} \tag{6.3}$$

The Gaussian function reads:

$$f_{\text{Gauss}}(x) = \exp\left(-0.5\left(\frac{x-\mu_{\text{Gauss}}}{\sigma_{\text{Gauss}}}\right)^2\right) \quad \text{with} \quad \begin{cases} \mu_{\text{Gauss}}\text{: mean value} \\ \sigma_{\text{Gauss}}\text{: width} \end{cases} \tag{6.4}$$

The latter probability density function is meant to describe the resolution effects.
Then the convolution of these two functions is realized through:

$$(f_{\text{Argus}} * f_{\text{Gauss}})(x) = \int f_{\text{Argus}}(x-y).f_{\text{Gauss}}(y).dy \quad \text{with} \quad \begin{cases} m_0 = m_{\Lambda_b^0} - m_{\pi^0} \simeq 5487 \text{ MeV} \\ \mu_{\text{Gauss}} = 0 \text{ MeV} \\ \sigma_{\text{Gauss}} = \sigma_{\Lambda_b^0} \end{cases} \tag{6.5}$$

The value of the threshold is fixed with respect to m_{π^0}, because this neutral meson is the lighter particle which can be "lost" during an event involving Λ_b^0 5-body decays for instance.

6.3 Combinatorial backgound

The combinatorial component is made of unrelated tracks which are mimicking the final states of interest.
It's often convenient to model these contributions with an exponential function. However, the signal spectrum seems very pure. A polynomial description might appear more relevant. Both shapes are examined in the following and the best description will be retained for the "charmless" fit model, which is not ion the scope of in this report.

The first approach is then to consider a Chebychev Polynomial[1] function of the first kind at the first order (i=1) as:

$$f_{\text{Cheby}}(x) = 1 + \sum_{i=1}^{n} a_i T_i(x) \quad \text{with} \quad \begin{cases} T_i\text{: Chebychev function of the first kind} \\ a_i\text{: coefficients.} \end{cases} \tag{6.6}$$

The second is to consider an exponential function as:

$$f_{\text{expo}}(x) = e^{\beta.x} \quad \text{with} \quad \beta \in \Re \tag{6.7}$$

[1] One can find in appendix B definitions of the first orders

6.4 Fit function

The individual probability density functions described previously are summed to form the total probability density function describing the full data spectrum. Note that two different probability density functions are built, one considering a polynomial combinatorics background and the second using an exponential one. The fit is done by using the unbinned extended likelihood minimization and determines the number of events of signal (N_{sig}), of background related to the convolution in equation 6.5 (N_{arg}) and of background related to equation 6.6 or equation 6.7 (N_{comb}).
In this way, the baseline fit function is given by the equation:

$$f_{\mathrm{fit}}(x) = N_{\mathrm{sig}}.(f_0.f_{\mathrm{CB1}}(x) + f_1.f_{\mathrm{CB2}}(x)) + N_{\mathrm{arg}}.(f_{\mathrm{Argus}} * f_{\mathrm{Gauss}})(x) + N_{\mathrm{comb}}.f^{\mathrm{expo}}_{\mathrm{Cheby}}(x) \tag{6.8}$$

Chapter 7

Fit Results

This fit model described in the previous chapter was initially built and improved with an earlier data set, containing around 7 times less statistics. Only results from the nominal data set are presented in the following.

7.1 Fit parameters

For all the different distributions of all decay modes, a simultaneous fit is performed and all free parameters shared, but the yields. The signal parameters describing the tails are fixed thanks to the relations (6.2). All the others[1] are allowed to get a value in a range fixed by educated values except the power p (6.3) which is arbitrarily fixed to 0, in order to have no divergence issue during the convolution. Values of the parameters are reported in the Table 7.1 below.

Fit parameters (data of 2012)	
Parameter name	Value
$\mu_{\Lambda_b^0}$	5624.0 ± 0.2 MeV
$\sigma_{\Lambda_b^0}$	17.0 ± 0.2 MeV
c	-0.9746 ± 0.4472
a_1	-0.9447 ± 0.0338
β	-0.0032 ± 0.0002

Table 7.1: Values of the parameters of the fit (data of 2012).

7.2 Yields per channel

The signal yields for each charmed decay of interest are gathered in Table 7.2 as given by the fit.

[1] in particular mean, slope of the Argus function, slope of combinatorics and number of events of the different categories

Signal Events (data of 2012)	
Decay mode	Yield
$\Lambda_b^0 \to \Lambda_c^+(\to p\pi^-\pi^+)\pi^-$	598 ± 25
$\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$	7539 ± 88
$\Lambda_b^0 \to \Lambda_c^+(\to pK^-K^+)\pi^-$	153 ± 13
$\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$	986 ± 34

Table 7.2: Number of events selected for each decay mode (fit to the data of 2012).

7.3 Invariant mass distributions and fit results displays

The figures 7.1 to 7.4 show the invariant mass distributions of the different reconstruction hypothesis with the result of the fit superimposed. All plots are reported in logarithmic scale.

Few remarks are in order:

- The amount of signal events is large and sufficient to consider these spectra as control channels for the decay modes $\Lambda_b^0 \to p\pi^+\pi^-\pi^+$ and $\Lambda_b^0 \to pK^+\pi^-\pi^+$;
- The fit model is well defined for the partially reconstructed background as well as for the signal ;
- The polynomial seems more relevant to describe the combinatorial component.
- The partially reconstructed background present a richer structure than the modelling we defined for it. Yet, a deep understanding of this part of the spectrum is not required for the education of the charmless fit model.

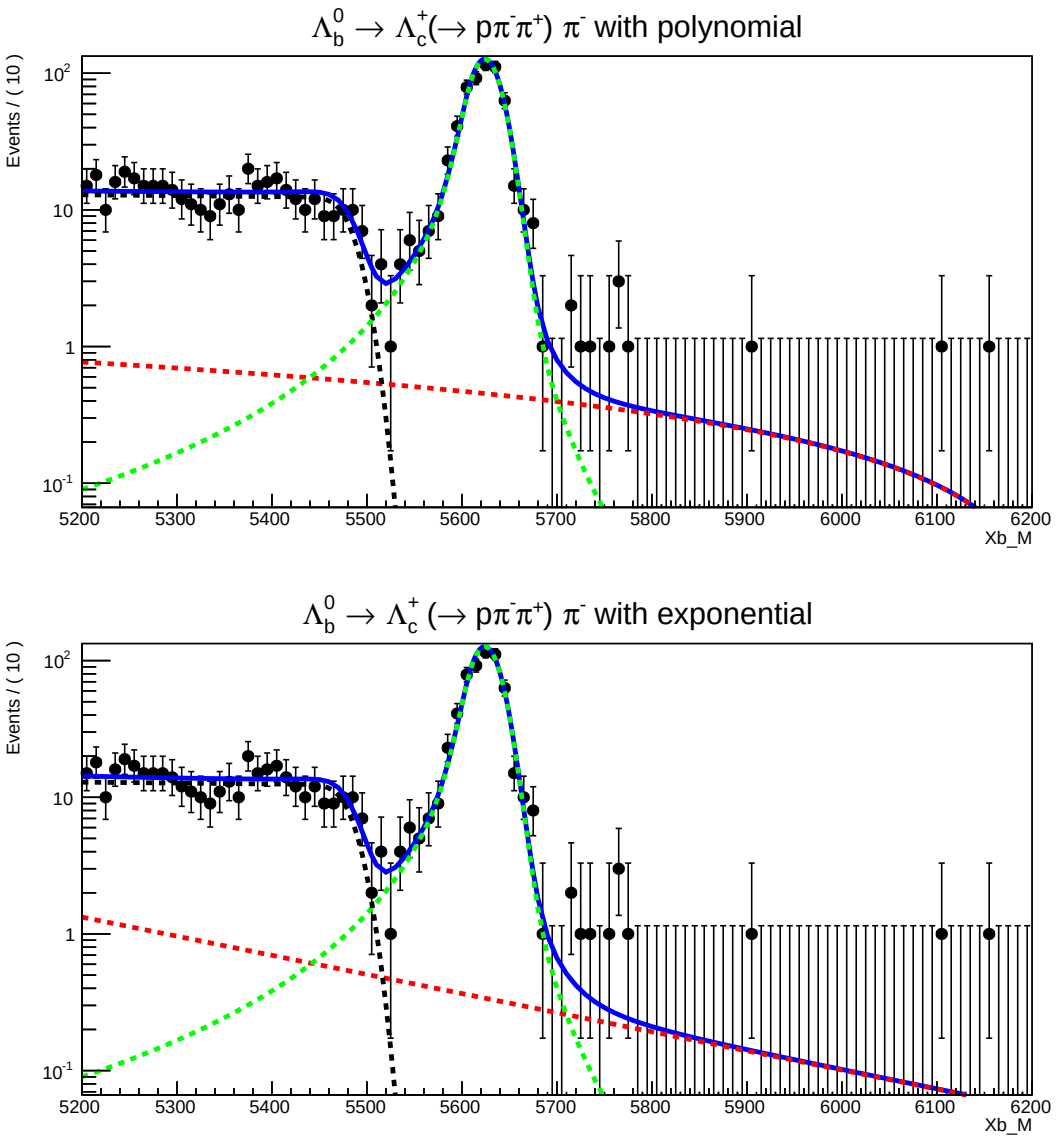


Figure 7.1: Fit of the invariant mass distribution for the decay $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^-\pi^+)\pi^-$ in logarithmic scale (data of 2012). The partially reconstructed background component is reported in black, the signal in green and the combinatorial component in red.

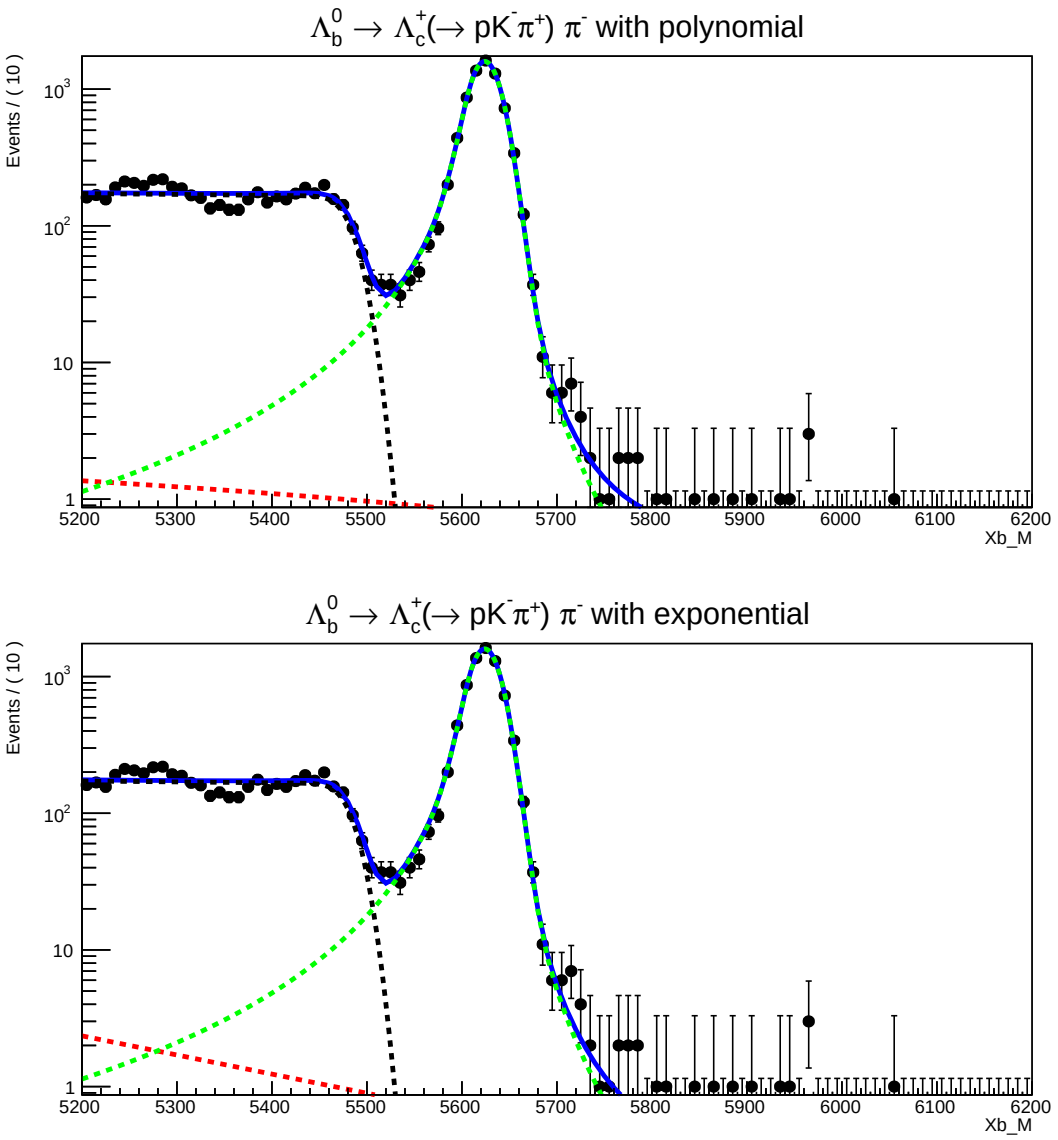


Figure 7.2: Fit of the invariant mass distribution for the decay $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ in logarithmic scale (data of 2012). The partially reconstructed background component is reported in black, the signal in green and the combinatorial component in red.

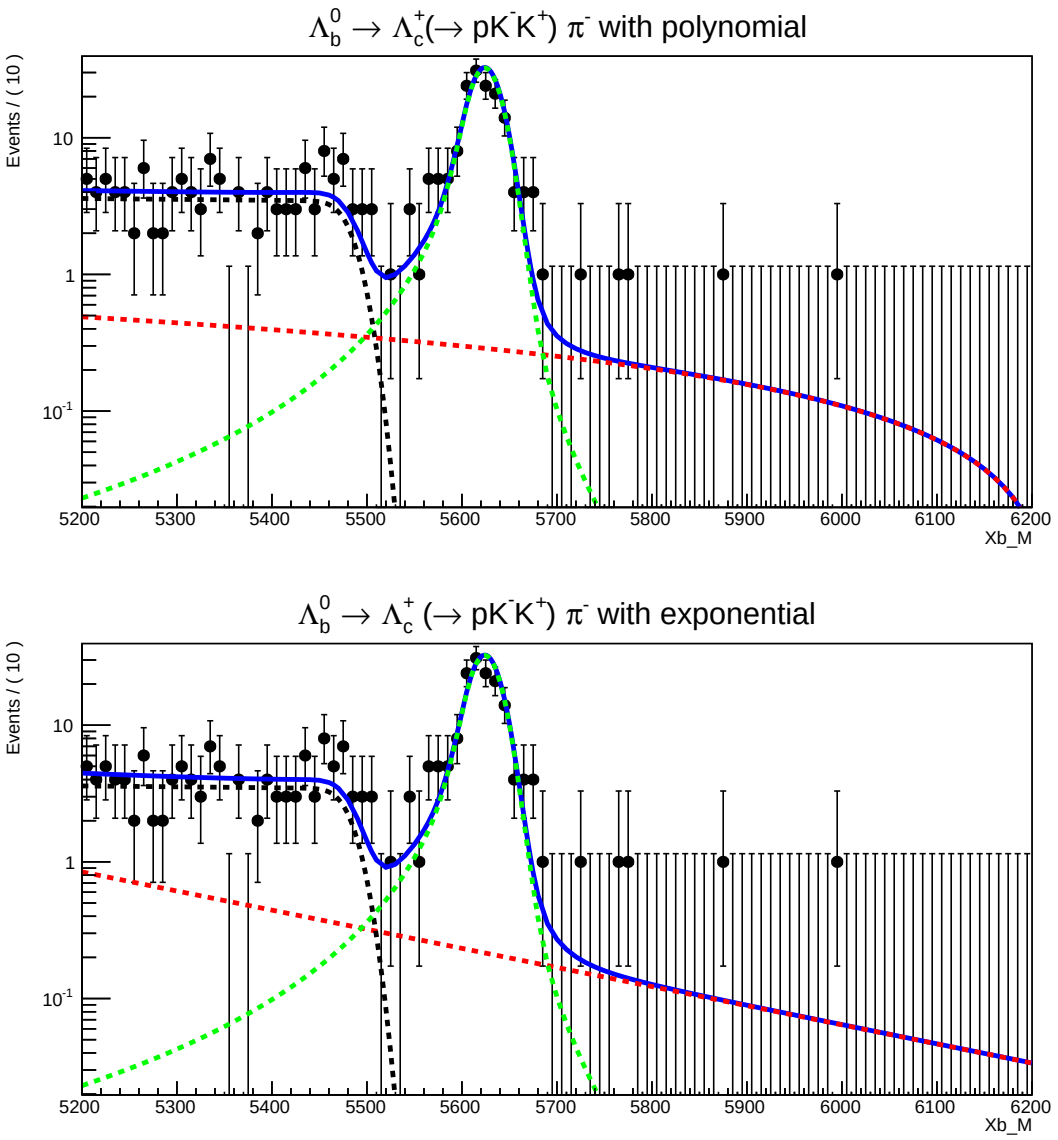


Figure 7.3: Fit of the invariant mass distribution for the decay $\Lambda_b^0 \to \Lambda_c^+(\to pK^-K^+)\pi^-$ in logarithmic scale (data of 2012). The partially reconstructed background component is reported in black, the signal in green and the combinatorial component in red.

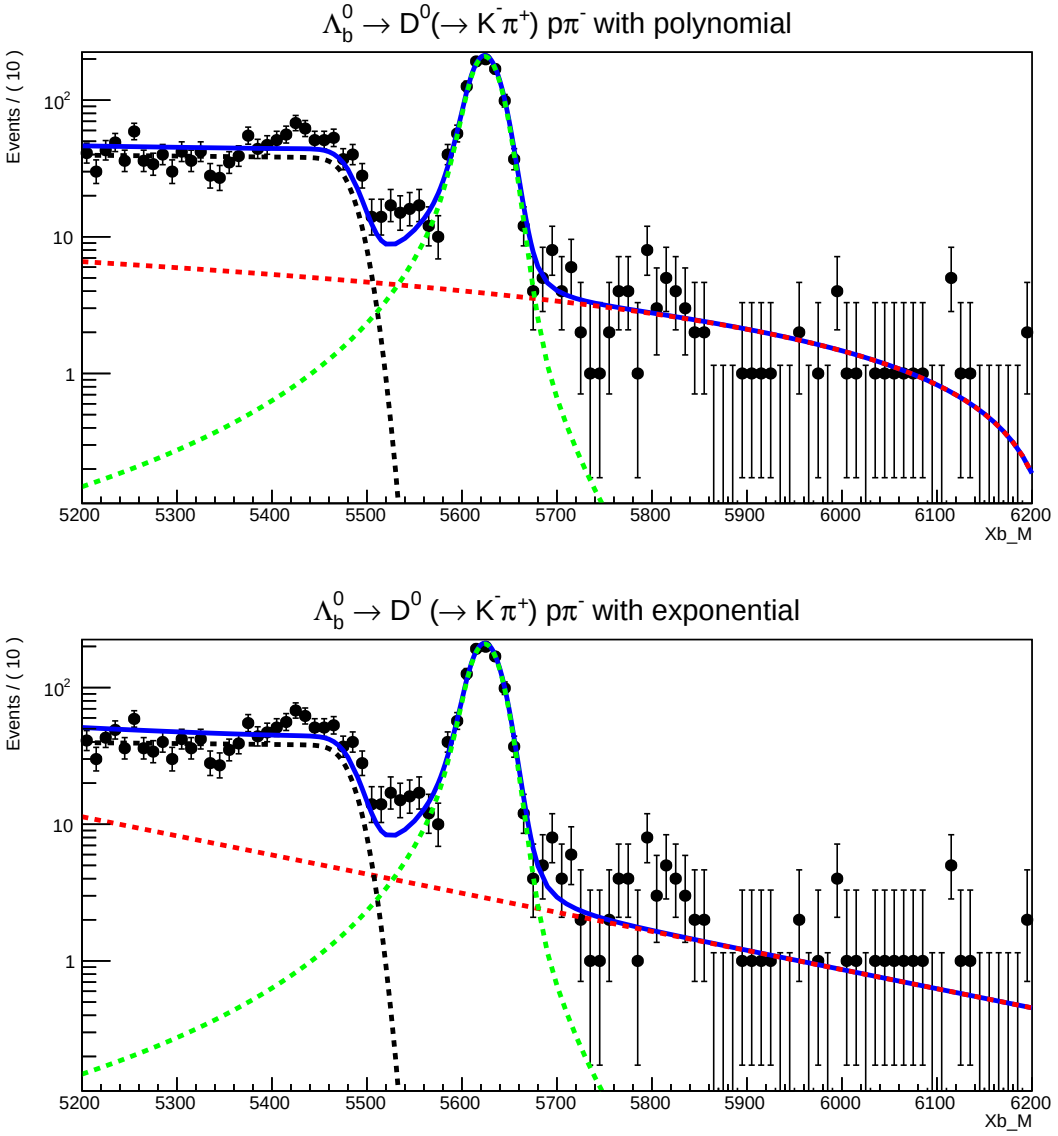


Figure 7.4: Fit of the invariant mass distribution for the decay $\Lambda_b^0 \to D^0(\to K^-\pi^+)p\pi^-$ in logarithmic scale (data of 2012). The partially reconstructed background component is reported in black, the signal in green and the combinatorial component in red.

Chapter 8

Estimation of B physics background yield present in the data

In the previous chapters, we have described and characterized the charmed decays serving as control channels for the study of CP violation in charmless decays. The purpose of this chapter is now to estimate the yields of background originatory from the decay of B mesons in 4-body where a final state hadron of the decay chain is misidentified as a proton, which are contained in the charmless sample after the selection procedure[1] . In the "charmless" analysis, the signal range is blind, so we have to extrapolate the B physics background contribution that we could find in the right-handed side-band of the spectrum. This work could allow to estimate the ratio between B physics background and combinatorics. This ratio can in turn be used to constrain the simultaneous fit of the charmless samples.

8.1 Estimation in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ spectrum

The right-handed side-band of this channel is made of events in which the invariant mass determined for Λ_b^0 is greater than 5685 MeV, which corresponds to 3 standard deviations above the nominal mass. For this set of events, the B physics background can be highlighted by looking at the $K^+\pi^-\pi^-\pi^-$ invariant mass distribution. In this case, for the dominant mode, the kaon produced in the decay chain is misidentified with a proton. These events are then contained in the data sample, but the invariant mass reconstructed under the assumption of a misidentification (kaon $\leftrightarrow$ proton) should be contained in a lower mass range, because of the difference of masses between those particles[2]. Several backgrounds have been imagined. In this analysis, two B decays are considered : $B \to K\pi\pi\pi$ (and partially reconstructed background) and $B \to a_1\pi$ (where $a_1 \to \pi\pi\pi$). These contributions have been modeled and the agreement between the fit model and the data is excellent. We then believe to have identified the main sources of B decay background. The yields, determined by this fit, are reported on the table 8.1. The invariant mass distribution as well as the fit are displayed on the figure 8.1.
Eventually for this process, we have $N_B^{\text{RHSB}} = 128 \pm 15$ and $N_{\text{comb}}^{\text{RHSB}} = 209 \pm 18$ which are the yields of B physics background and of combinatorics present in the right-handed side-band, respectively. The ratio R between these quantities for the full mass range we are interested in, can be estimated by using integrals of each of these contributions (extrapolated in the blind signal range) thanks to the equation 8.1.

[1] One can find in appendix C same yield estimation with another selection on the proton PID variable
[2] $m_p - m_{K^\pm} \simeq 444.6$ MeV

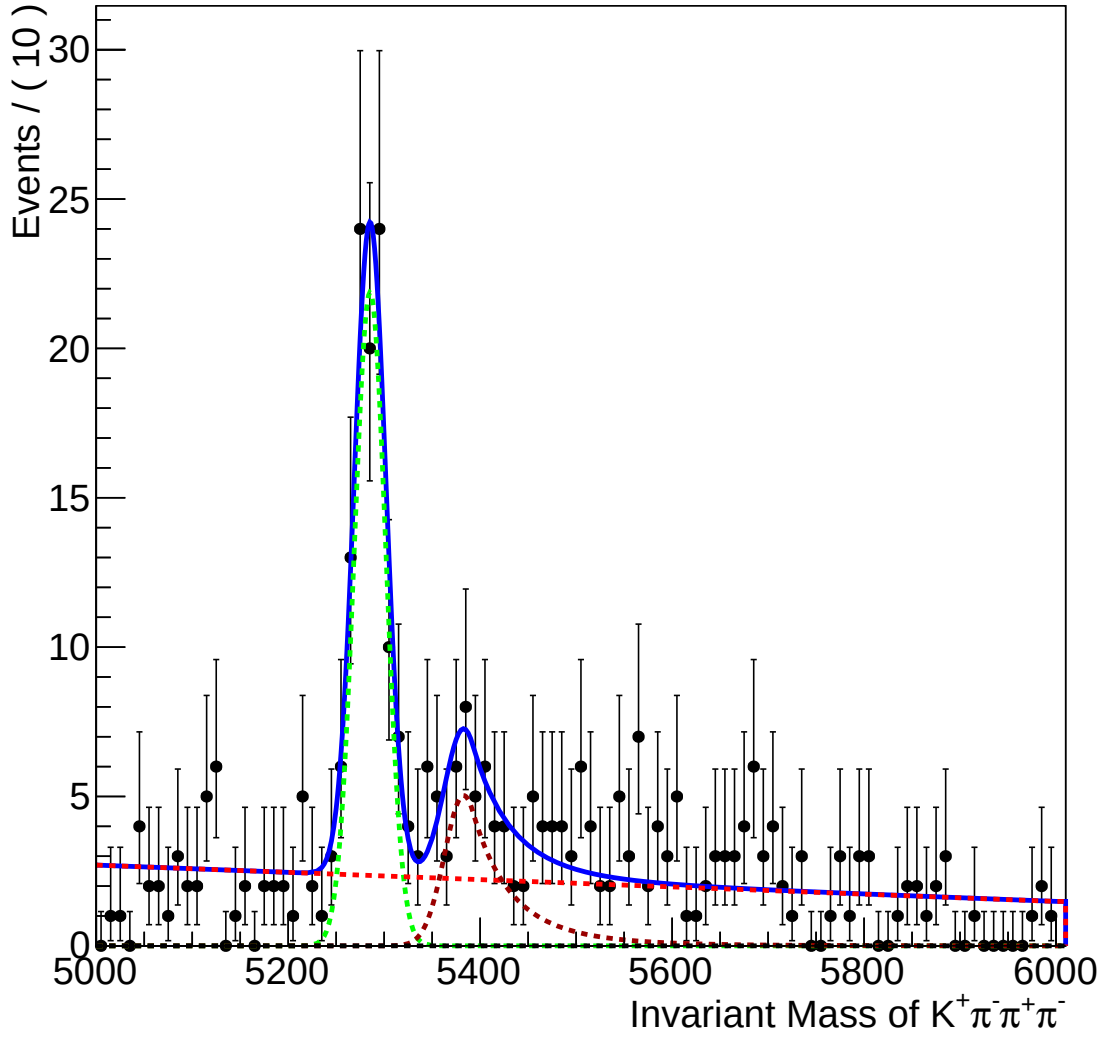


Figure 8.1: Fit of $\Lambda_b^0 \to p\pi\pi\pi$ candidates in the right-handed side-band, reconstructed as $K\pi\pi\pi$ for 2011. $B \to K\pi\pi\pi$ contribution is reported in green, $B \to a_1(\to \pi\pi\pi)\pi$ one in brown, the partially reconstructed background in black and the combinatorial component in red. The latter are thought to be $p\pi\pi\pi$ combinatorial background.

$$R = \frac{N_{\text{comb}}^{\text{spectra}}}{N_{\text{B}}^{\text{spectra}}} = \frac{I_{\text{B}}}{I_{\text{comb}}} * \frac{N_{\text{comb}}^{\text{RHSB}}}{N_{\text{B}}^{\text{RHSB}}} \tag{8.1}$$

Thanks to the integral reported in the table 8.2, which are evaluated from the probability density function used in the charmless fit, we get $R = 0.55 \pm 0.07$.

Process	Number of events
$B \to K\pi\pi\pi$	90 ± 11
$B \to a_1\pi$	38 ± 10
Partially reconstructed Background	0 ± 3
Combinatorics	209 ± 18

Table 8.1: Yields of background processes determined by the fit in the right-handed side-band of the spectra $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ with $K \leftrightarrow p$.

Background sources	Integral
B physics	0.151
Combinatorics	0.454

Table 8.2: Fraction of background sources in the right-handed side band for $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$

8.2 Estimation in the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ spectrum

For this spectrum, the right-handed side-band is defined from $M_{X_b} > 5840$ Mev/c^2, in order to not unblind Ξ_b^0 signal range. The very same work is done with the spectrum $\Lambda_b^0 \to pK\pi\pi$ by considering the misidentification (K $\leftrightarrow$ p), but also an additional misidentification $\pi \leftrightarrow K$ or $K \leftrightarrow \pi$. The reference spectrum is chosen to be $KK\pi\pi$. Yields determined by the fit are reported on the table 8.3, and the invariant mass distribution fitted is visible on the figure 8.2. We have identified three sources of B meson decays : $B \to KK\pi\pi$ (single misidentification $p \leftrightarrow K$), $B \to KKK\pi$ and $B \to \pi K\pi\pi$ (one additional misidentification $\pi \leftrightarrow K$). Again, the quality of that fit suggests that these background sources are basically identified.

Process	Number of events
$B \to KK\pi\pi$	64 ± 10
$B \to KKK\pi$	46 ± 13
$B \to \pi K\pi\pi$	49 ± 11
Partially reconstructed Background	1 ± 15
Combinatorics	173 ± 20

Table 8.3: Yields of background processes determined by the fit in right-handed side-band of the spectra $\Lambda_b^0 \to pK^-\pi^+\pi^-$ with $K \leftrightarrow p$.

Thanks to the integral reported in the table 8.4 we get for this spectrum $R = 1.10 \pm 0.15$. These ratios provide a benchmark result which should be used to judge the quality of the charmless fit.

Background sources	Integral
B physics	0.281
Combinatorics	0.280

Table 8.4: Fraction of background sources in the right-handed side band for $\Lambda_b^0 \to pK^-\pi^+\pi^-$

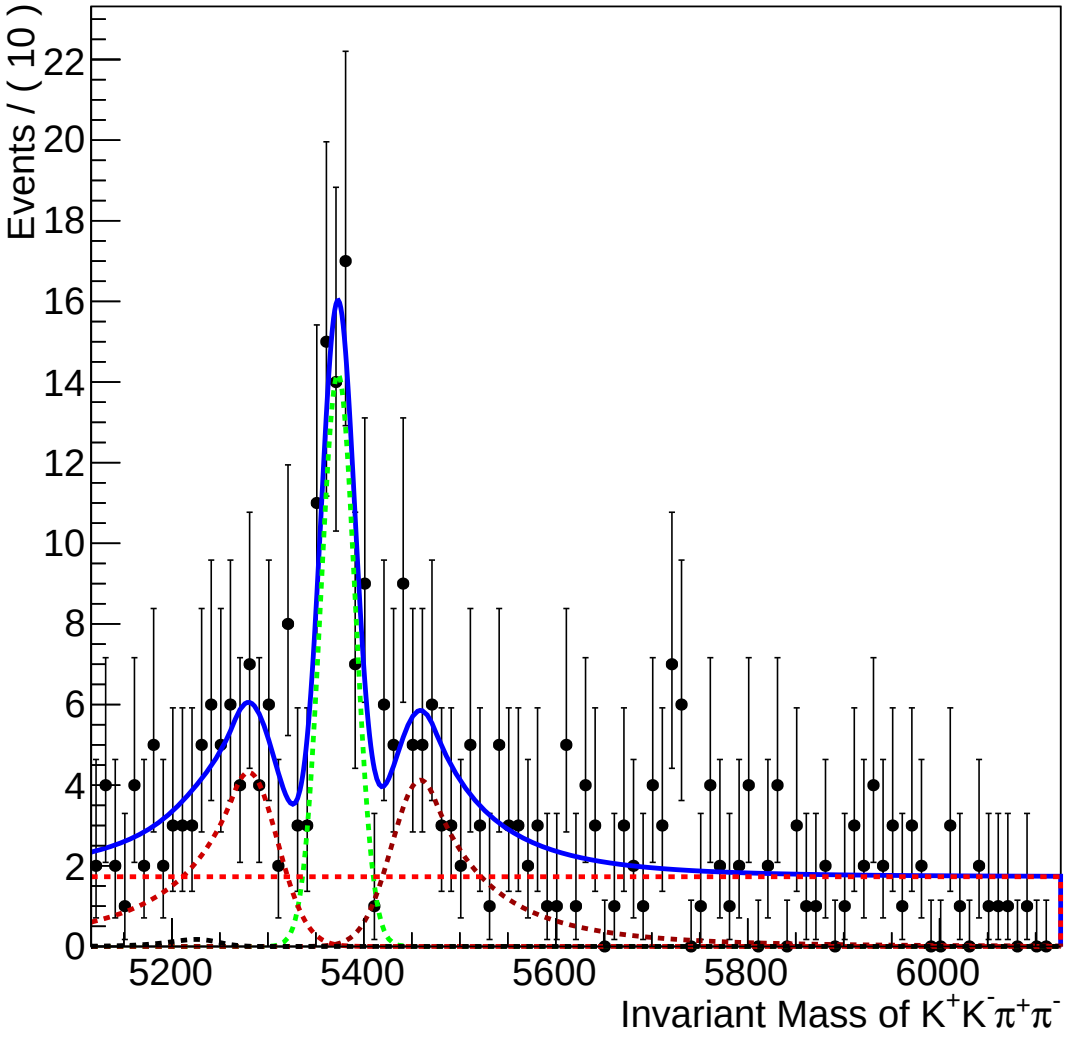


Figure 8.2: Fit of $\Lambda_b^0 \to pK\pi\pi$ reconstructed as $KK\pi\pi$ for 2012. $B \to KK\pi\pi$ contribution is reported in green, $B \to\to KKK\pi$ and $B \to \pi K\pi\pi$ ones in brown, the partially reconstructed background in black and the combinatorial component in red.

Conclusion

CP-asymmetries, experimentally observed in mesons decays and mixing, are currently remarkably described by the single CP-violating phase of the Standard Model. The CP violation in baryon decays is unobserved to date and would provide a complementary and invaluable input in this landscape. The charmless b-baryon 4-body decays might be a relevant place to perform this very first observation. My master internship gives a contribution to this quest.
In order to select events of interest, a multivariate analysis has been performed, and gives an efficient separation between signal and combinatorial background. Moreover, CP-asymmetries measurement require to characterize detection efficiencies as well as asymmetries of production. Events containing same final states of interest, which are reached via $b \to c$ transitions, can be used as control channels. In the analysis presented in this work, a fit model is built and shows satisfactory results to describe these charmed decays. Eventually, the knowledge of the background sources present in the spectra of interest must be mastered. We conducted a dedicated analysis to understand their origin and believe that they are now under control.

Part III

APPENDICES

Appendix A

Boosted Decision Tree (BDT)

Let's consider we want to separate signal and background events and we have Monte-Carlo samples for each. We first have to *train* the decision tree and then, *test* it.

A.1 Training

For each event to separate, we can find usefull seperation variables to identify properties linked to the signal (vertexing variables for example). We first have to order events with respect to values of each variable and see what happens, in term of signal-background separation, if we split in two parts the training sample for a given value of a given variable. Then, we pick the value which gives the best separation and do the same with another variable. In this way we have two *leaves* for which we do the same process again. The question is when to stop this process: we could fix a number of branches for instance or stop when each leaf is pure enough in terms of background or signal events present in this leaf. For the latter case, a leaf is called signal or background leaf with respect to a criterion called *Gini* index defined as:

$$Gini = (\Sigma_{i=1}^{n} W_i)\, P(1-P) \quad \text{with} \begin{cases} W_i = \text{ Weight of the } i^{th} \text{ event.} \\ n\text{: number of events in the leaf.} \\ P = \dfrac{\Sigma_s W_s}{\Sigma_s W_s + \Sigma_b W_b}\text{: leaf purity.} \end{cases}$$

The criterion chosen is to minimize $Gini_{\text{left-son}} + Gini_{\text{right-son}}$ and eventually, we hand up with a decision tree.
Boosting a decision tree is dealing with misidentified events[1]. If an event is misidentified, we weight (boost) it and start again the training. In this way, one can build many trees. It exists different ways to boost decision trees, depending how events are weighted. In our case, an Adaptive Boost is used[11].

A.2 Test

Once the training is done, we apply the best seperation procedure found to another sample called test sample. Training and test samples are built randomly, it means that results with one of

[1] background events in a signal leaf or vice-versa

them have to be very similar with those done with the other sample. One possible test to verify the agreement between training and test is the Kolmogorov-Smirnov one.

A.3 Overtraining

One of the main error that we have to fight against is the overtraining. It happens when the training step is actually dealing with statistical fluctuations to seperate signal event and background one. It is mandatory, to get the best seperation possible, to not have any overtraining.

Appendix B

Chebychev polynomials

As described in the section 6, we use Chebychev polynomials of the first kind at the first order to model combinatorial component of the fit. These polynomial functions are defined as:

$$\cos(n\theta) = T_n \cos(\theta) \tag{B.1}$$

where:

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \quad \text{with:} \quad \begin{cases} T_0 = 1 \\ T_1 = x \end{cases} \tag{B.2}$$

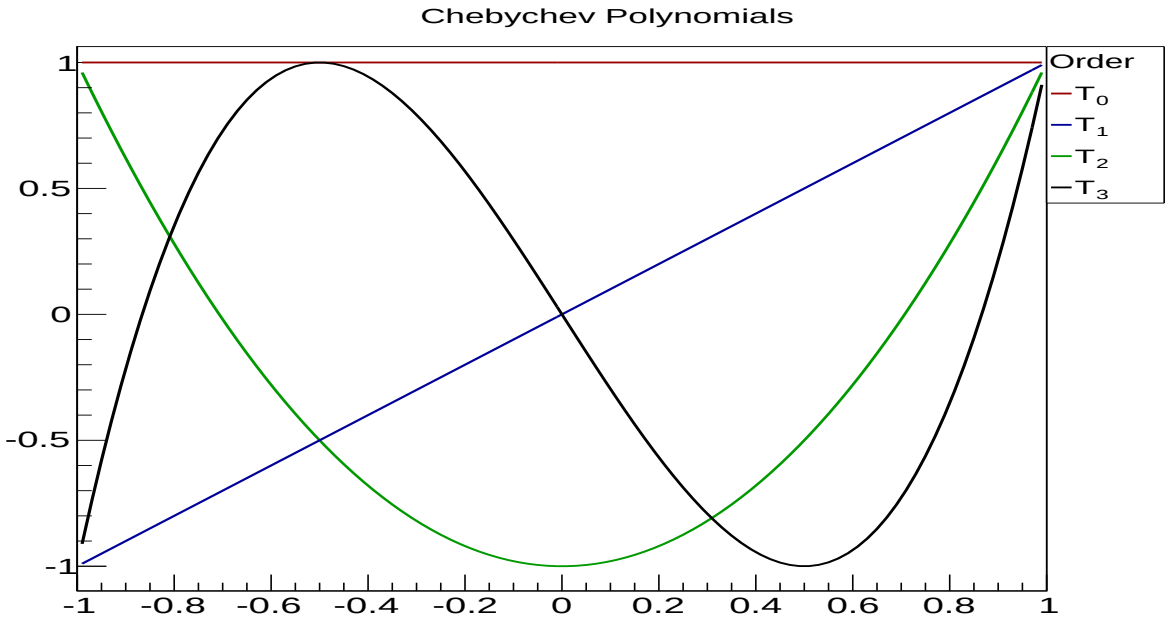


Figure B.1: Chebychev polynomials at order $n = 0$, $n = 1$, $n = 2$ and $n = 3$

Appendix C

B physics background estimation with other cut on PID variable of the proton.

In this appendix are presented similar results as ones displayed in chapter 8 but with a softer selection on the proton PID variable. The different contributions have been initially modeled with this selection.

Process	Number of events
$B \to K\pi\pi\pi$	1072 ± 36
$B \to a_1\pi$	166 ± 21
Partially reconstructed Background	326 ± 29
Combinatorics	854 ± 52

Table C.1: Yields of background processes determine by the fit in right-handed side-band of the spectra $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ with $K \leftrightarrow p$

Process	Number of events
$B \to KK\pi\pi$	653 ± 28
$B \to KKK\pi$	440 ± 31
$B \to \pi K\pi\pi$	274 ± 22
Partially reconstructed Background	166 ± 24
Combinatorics	285 ± 28

Table C.2: Yields of background processes determine by the fit in right-handed side-band of the spectra $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ with $K \leftrightarrow p$

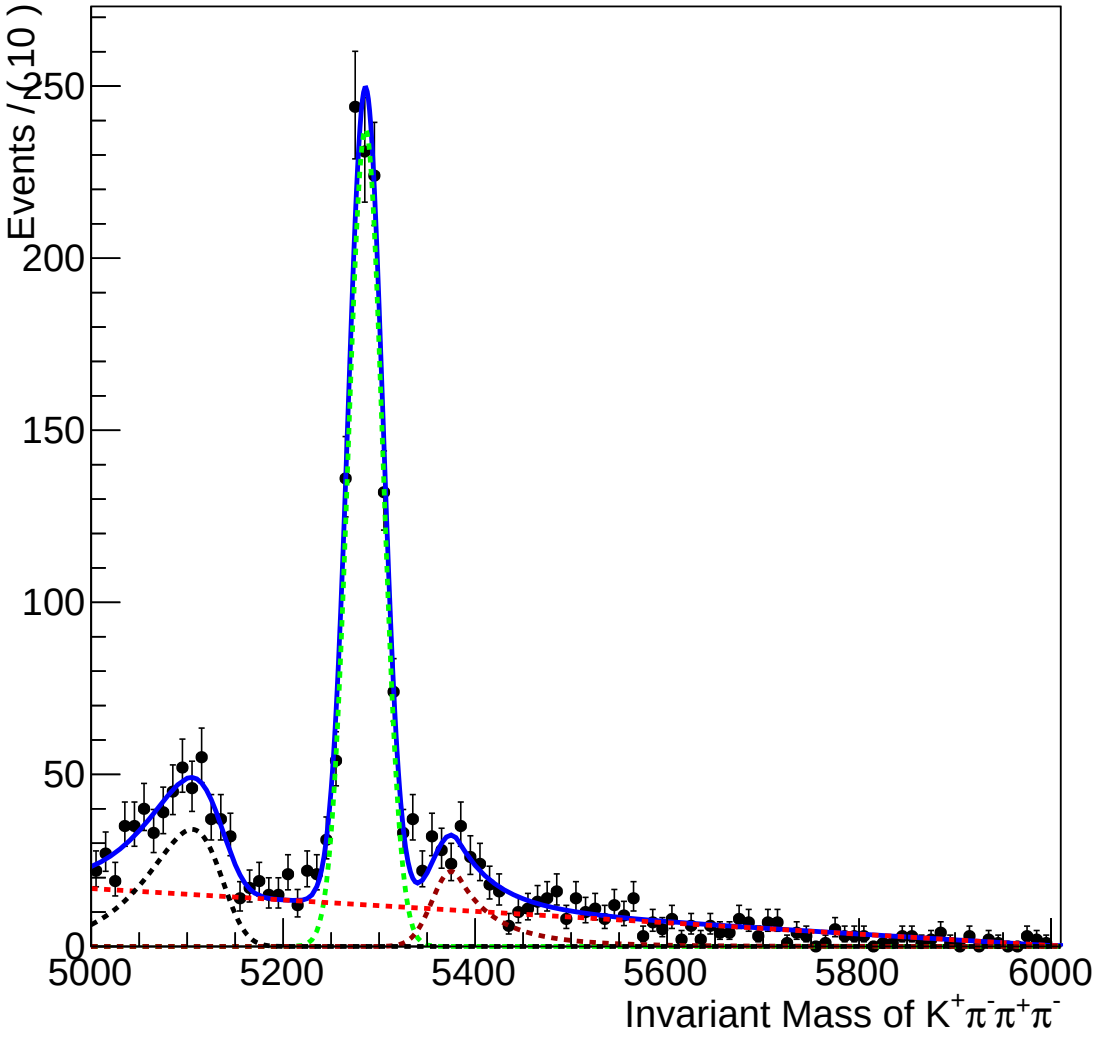


Figure C.1: Fit of $\Lambda_b^0 \to p\pi\pi\pi$ reconstructed as $K\pi\pi\pi$ for 2011. $B \to K\pi\pi\pi$ contribution is reported in green, $B \to a_1(\to \pi\pi\pi)\pi$ one in brown, the partially reconstructed background in black and the combinatorial component in red.

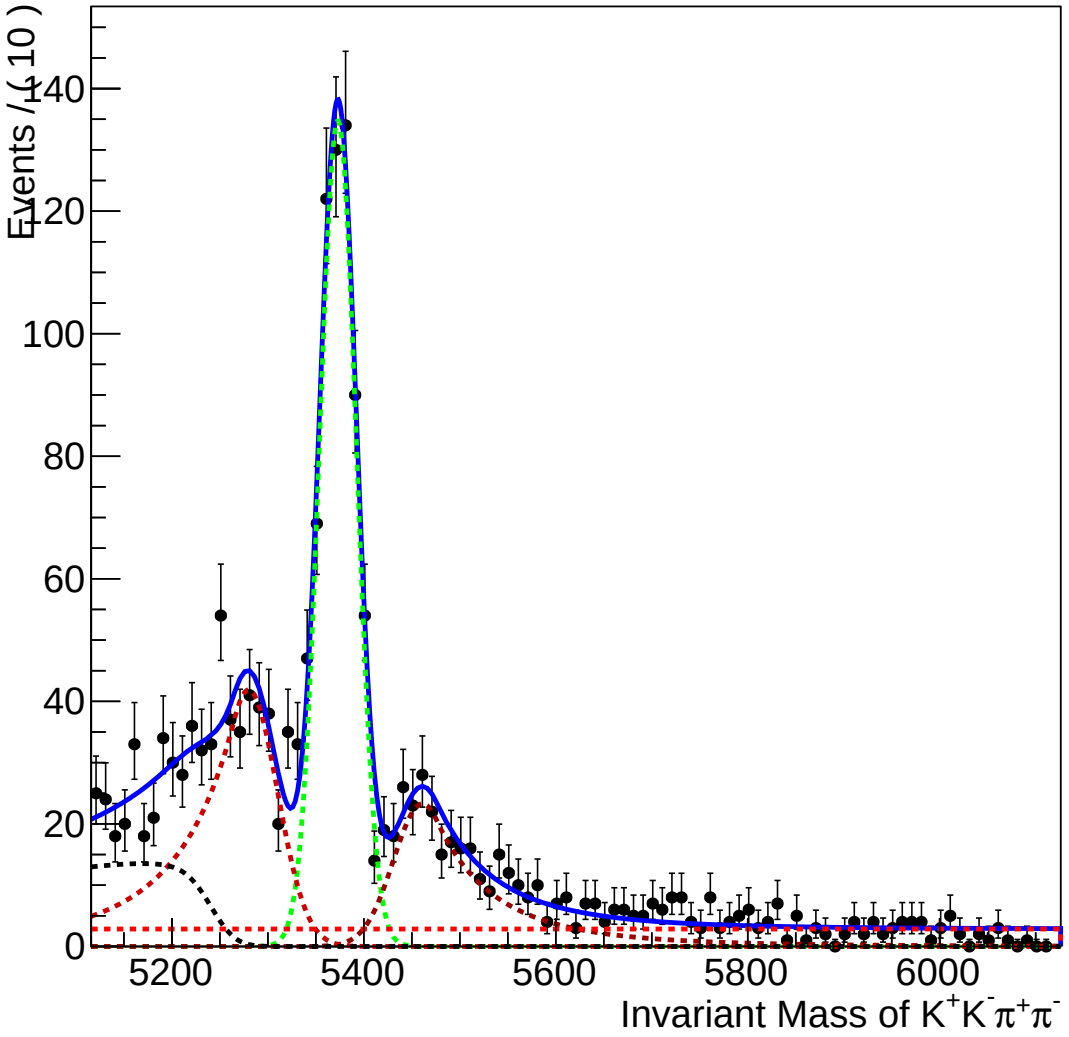


Figure C.2: Fit of $\Lambda_b^0 \to pK\pi\pi$ reconstructed as $KK\pi\pi$ for 2012. $B \to KK\pi\pi$ contribution is reported in green, $B \to\to KKK\pi$ and $B \to \pi K\pi\pi$ ones in brown, the partially reconstructed background in black and the combinatorial component in red.

Bibliography

[1] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[2] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.

[3] A. Salam. Weak and electromagnetic interactions. 1969.

[4] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[5] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[6] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[7] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964.

[8] Particle data group, 2012. www.pdg.lbl.gov.

[9] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1 – 29, 2012.

[10] D.Kirby W.Verkerke. *RooFit Users Manual v2.07.*

[11] Byron P. Roe, Ji Zhu Hai-Jun Yang, Ion Stancu Yong Liu, and Gordon McGregor. Boosted decision trees as an alternative to artificial neural networks for particle identificaton. Nov 2004.

[12] Stéphane Monteil. *Mesures de précision électrofaibles.* Habilitation à diriger des recherches, Université Blaise Pascal, December 2009.

Abstract

CP-asymmetries, experimentally observed in mesons decays and mixing, are currently remarkably described by the single CP-violating phase of the Standard Model. The CP violation in baryon decays is unobserved to date and would provide a complementary and invaluable input in this landscape. The charmless b-baryon 4-body decays might be a relevant place to perform this very first observation. My master internship gives a contribution to this quest.
In order to select events of interest, a multivariate analysis has been performed, and gives an efficient separation between signal and combinatorial background. Moreover, CP-asymmetries measurement require to characterize detection efficiencies as well as asymmetries of production. Events containing same final states of interest, which are reached via $b \to c$ transitions, can be used as control channels. In the analysis presented in this work, a fit model is built and shows satisfactory results to describe these charmed decays. Eventually, the knowledge of the background sources present in the spectra of interest must be mastered. We conducted a dedicated analysis to understand their origin and believe that they are now under control.

Keywords: ElectroWeak Standard Model, CP violation, Four-body decay

Résumé

Les asymétries CP, expérimentalement observées dans les désintégrations et les mélanges de mésons, sont décrits avec succès par une phase unique dans le Modèle Standard. La violation de la symétrie CP dans les désintégrations baryoniques reste inobservée, mais apporterait une avancée incontestable dans ce domaine. Les désintégrations en quatre corps sans charme de baryons beaux peut être le lieu pertinent pour une toute première observation. Ce rapport apporte une contribution à cette recherche.
Dans le but de sélectionner les événements d'intérêt, une analyse multivariable est réalisée, et mène à une séparation efficace entre le signal et le bruit de fond de combinatoire. La mesure des asymétries CP nécessite également la caractérisation des efficacités de détection et des asymétries de production. Les événements qui contiennent les mêmes états finals d'intérêt, opérant via des transitions $b \to c$, peuvent être utilisés comme canaux de contrôle. Dans l'analyse présentée dans ce document, un modèle d'ajustement des distributions de masse invariante de ces canaux, est construit et présente de satisfaisants résultats quant à la description de ces désintégrations charmées. Finalement, la connaissance des sources de bruit de fond présentes dans les spectres d'intérêts doit être maitrisée. Une analyse est menée pour comprendre leurs origines et nous laisse penser qu'elles sont à présent sous contrôle.

Mots clefs: Modèle Standard Electrofaible, Violation CP, Désintégration en quatre corps